

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 20:42:00 ; Search time 655.99 Seconds
(without alignment)
5125.708 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999
Sequence: 1 MKMKKQEVVVTGPTKGDV.....KLCTSKPTSQVPLAQRITFF 568

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=x1h
-Q=/cgr2.1/USPTO.spool/US09904994/runat.11022005.145646.25526/app.query.fasta_1.1102
-DB=N GeneSeq.16Dec04 -QPM=fastcap -SUPFIX=ring -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=p2n -NDBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09904994.QCGN.1.1.468@runat.11022005.145646.25526 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N_GeneSeq.16Dec04: *
1: geneSeq1980s: *
2: geneSeq1990s: *
3: geneSeq2000s: *
4: geneSeq2001as: *
5: geneSeq2001bs: *
6: geneSeq2002as: *
7: geneSeq2002bs: *
8: geneSeq2003as: *
9: geneSeq2003bs: *
10: geneSeq2003cs: *
11: geneSeq2003ds: *
12: geneSeq2004as: *
13: geneSeq2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2999	100.0	2883	7	ADJ58237 Urease su
2	2973	99.1	2452	7	ADJ58249 UreaseXY
3	2960	98.7	2405	7	ADJ58240 UreaseXY
4	2953	98.5	2407	7	ADJ58246 UreaseXY
5	2588	86.3	2183	7	ADJ58243 UreaseXY

6	2326	77.6	1710	12	ADQ37850	Adq37850 H. bizzo2
7	2326	77.6	8407	12	ADQ37847	Adq37847 H. bizzo2
8	2297	76.6	1710	2	AAQ04329	AAQ04329 Probe for
9	2297	76.6	1710	8	ACA34557	ACA34557 Prokaryot
10	2297	76.6	1815	2	AAK14230	AAK14230 H. pylori
11	2297	76.6	2735	2	AAT44351	AAT44351 H. pylori
12	2291	76.4	1710	9	AAI60579	AAI60579 Helicobac
13	2286	76.2	1717	7	ABA00817	ABA00817 H. pylori
14	2258	75.3	1719	8	ABA00816	ABA00816 H. felis
15	2250	75.0	2767	2	AAQ12485	AAQ12485 DNA encod
16	2244	74.8	2619	2	AAQ09180	AAQ09180 Helicobac
17	2244	74.8	2619	2	AAT45680	AAT45680 H. felis
18	2241	74.7	2619	2	AAQ075319	AAQ075319 Urease ur
19	2233	74.5	2385	2	AAV62460	AAV62460 Helicobac
20	1922	64.1	1704	8	ACA23584	ACA23584 Prokaryot
21	1914	63.8	1698	8	ACA45232	ACA45232 Prokaryot
22	1907	63.6	1704	8	ACA25712	ACA25712 Prokaryot
23	1897	63.3	1704	8	ACA26184	ACA26184 Prokaryot
24	1832	61.1	1701	8	ACA42689	ACA42689 Prokaryot
25	1832	61.1	3919	13	ADT05424	ADT05424 Haemophil
26	1832	61.1	85814	13	ADT05644	ADT05644 Haemophil
27	1828	61.0	1701	11	ABD11712	ABD11712 Pseudomon
28	1826	60.9	1719	8	ACA34156	ACA34156 Prokaryot
29	1826	60.9	110000	2	AAT42063_05	Continuation (6 of
30	1825	60.9	1701	8	ACA35953	ACA35953 Prokaryot
31	1821	60.7	1878	11	ACH96250	ACH96250 Klebsiell
32	1818.5	60.6	1710	10	ADF01572	ADF01572 Bacteri
33	1818.5	60.6	2341	8	ACA45080	ACA45080 Prokaryot
34	1818	60.6	1704	9	ADA30499	ADA30499 DNA encod
35	1817	60.6	2400	2	AAV19002	AAV19002 Klebsiell
36	1817	60.6	4768	3	AAZ92097	AAZ92097 Klebsiell
37	1815	60.5	1698	8	ACA20910	ACA20910 Prokaryot
38	1809	60.3	1707	8	ACA32769	ACA32769 Prokaryot
39	1809	60.3	1707	8	ACA32790	ACA32790 Prokaryot
40	1809	60.3	86248	10	ADC00087	ADC00087 Enterococ
41	1809	60.3	87563	9	ACD19044	ACD19044 E. coli 0
42	1808	60.3	8729	9	ACD19244	ACD19244 E. coli 0
43	1797	59.9	1701	8	ACA44198	ACA44198 Prokaryot
44	1789.5	59.7	1716	4	AAH52362	AAH52362 S. epider
45	1789.5	59.7	2439	8	ACA46520	ACA46520 Prokaryot

ALIGNMENTS

RESULT 1	
ADJ58237	
ID	ADJ58237 standard; DNA; 2883 BP.
XX	
AC	ADJ58237;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Urease subunit polypeptide complex encoding sequence.
XX	
KW	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX	
OS	Helicobacter felis.
XX	
FT	Key
FT	CD8
FT	Location/Qualifiers
FT	206..886
FT	/*tag= a
FT	/product= "urease polypeptide complex"
FT	897..2603
FT	/*tag= b
FT	/product= "urease polypeptide complex"
XX	
PN	EP1176192-A2.
XX	
PD	30-JAN-2002.
XX	
PF	11-JUL-2001; 2001BP-00202666.
XX	
PR	17-JUL-2000; 2000BP-00202565.

QY 561 LeuAlaGlnArgIYThrPhePhe 568
 DB 2577 CTAGCCAGCGCTACACTTCTTC 2600

RESULT 2
 ID ADJ58249
 ADJ58249 standard; DNA; 2452 BP.
 AC ADJ58249;
 DT 06-MAY-2004 (first entry)
 XX ureaseX subunit encoding sequence #4.
 DE ureaseX subunit encoding sequence #4.
 KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
 XX Helicobacter felis.
 OS Helicobacter felis.
 FH Key Location/Qualifiers
 FT CDS 48..728
 FT /*tag= a
 FT /product= "urease protein"
 FT CDS 739..2445
 FT /*tag= b
 FT /product= "urease protein"

EP1176192-A2.
 PD 30-JAN-2002.
 PF 11-JUL-2001; 2001EP-00202666.
 XX 17-JUL-2000; 2000EP-00202565.
 PR (ALKU) AKZO NOBEL NV.
 PI Kusters JG, Cattoij G;
 XX WPI; 2002-124384/17.
 DR P-PSDB; ADJ58250, ADJ58251.
 XX

PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
 PT the diagnosis of Helicobacter felis infections and in the preparation of
 PT vaccines.
 PS Disclosure; SEQ ID NO 13; 76bp; English.
 XX

CC The present invention relates to a novel Helicobacter felis urease X and
 CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
 CC used in the manufacture of vaccines against Helicobacter felis infections
 CC and in diagnostic tests to detect antibodies against Helicobacter felis.
 CC Helicobacter felis is difficult to grow so it is more convenient to use
 CC the expression products of the genes encoding the urease X and Y subunits
 CC in the manufacture of vaccines. The present sequence represents an
 CC ureaseX subunit of the invention.

XX
 SQ Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,566-261 Length: 2452
 Score: 2973.00 Matches: 562
 Percent Similarity: 99.65% Conservative: 4
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 99.13% Indels: 0
 DB: 7 Gaps: 0

US-09-904-994B-3 (1-568) x ADJ58249 (1-2452)

QY 1 MetLysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
 DB 739 ATGAAATGAAAAAACAAGATATGTAATACCTACGACCAACCAACGCGATTAAGTG 798

QY 21 ArgLeuGlnYAspThrAspLeuTrpAlaGluValGlnHisAspTyrThrThrTyrGlyGlu 40
 DB 799 CGCTTAGAGATACGATCTTTGGCAGAAAGTAAAGATGCTATACCACTTAGCGCAA 858

QY 41 GLeuLysPheGlyAlaGlyLysThrLLeaGslGluGlyMetGlyGlnSerAsnSerPro 60
 DB 859 GAACCTCAATTCGGGTGACGTAAGAAAATCTATCCGTAGGGGTATGGGTCAAGCAATAGCCCA 918

QY 61 AspGluAsnThrLeuAspLeuValLLeaThrAsnAlaMetLLeaAspTyrThrGlyLLe 80
 DB 919 GATGAAAAACCTTAGATTTAGTATGATCACCAACCGATGATTAATGATCAACCGGGATT 978

QY 81 TyrLysAlaAspLLeaGlyLLeaValAsnGlyLysLLeaGlyLLeaGlyLysAlaGlyAsn 100
 DB 979 TACAAAGCCGACATTGGCATTAATAAATGCAAAATGCAATGCGATTTGCAAGCGAAAC 1038

QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeu 120
 DB 1039 AAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCCGTGGGTGGGCAAGAAACATA 1098

QY 121 AlaGlyGluGlyMetLLeaThrAlaGlyGlyLLeaAspSerHisThrHisPheLeuSer 140
 DB 1099 GCAGGGGAAGGATGATTAATTAACCGCTGGGGGATCGATTACACACCACTTCTCTCT 1158

QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrMetPheGlyGlyGlyThr 160
 DB 1159 CCACAACAATTCCTTACCGCTAGCCATGCAATGGCCTTCAACAAATGTTGGGCGGCACA 1218

QY 161 GlyProValAspGlyLysAsnAlaThrThrLLeaThrProGlyLysTPAsnLLeaHisArg 180
 DB 1219 GGCCCGTAGATGCGCAAGATGCGATACCATCACTCCGGCAATGGAACTTGCACCGC 1278

QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
 DB 1279 ATGTTCGCGCGCAGCAAGATATTCTATGATGTGGCTTTTGGCCAAAGGCAATAGC 1338

QY 201 SerSerLysLysGlnLeuValGlnGlnValGlnAlaGlyLLeaGlyPheLysLeuHis 220
 DB 1339 TCTAGTAAACCACTGTGTAAGCAAGTAGAACGGGCGGATGTGTTTAAATGTGAT 1398

QY 221 GlnAspTrpGlyLysThrProSerAlaLLeaAspIscLysLeuSerValAlaAspGluTyr 240
 DB 1399 GAACACTGGGGCAACAATCCAAAGCGCATGATGATACGCTTGAACGTGACGATGATAC 1458

QY 241 AspValGlnValCysLLeaThrAspThrValAsnGluAlaGlyLysValAspAspThr 260
 DB 1459 GATGTGCAAGTTGTATATACACACCGATACGTCATATGAGGAGGATGTATGATGACACC 1518

QY 261 LeuAsnAlaMetAsnGlyArgAlaLLeaAlaThrHisLLeaGluGlyAlaGlyGly 280
 DB 1519 CTAAATCAATGAACGGGCGCGCATCGCTATGCAACATGAGGAGCGGGTGAAGA 1578

QY 281 HisSerProAspValLLeaThrMetAlaGlyGluLeuAsnLLeuProSerSerThrThr 300
 DB 1579 CACTCACTGATGTATACCATGATGCGAGGAGTGAATTTCACTCTCTCCACAACC 1638

QY 301 ProThrLLeaProTyrThrLLeaThrValAlaGluHisLeuAspMetLeuMetThrCys 320
 DB 1639 CCTACTATCCCTTATACATTAATATACGGTTGCAAGAACATTAACATGCTTATATACCTGC 1698

QY 321 HisHisLeuAspLysArgLLeaGluLysLeuGlnPheSerGlnSerArgLLeaArgPro 340
 DB 1699 CACCACTGATGAACCGATCCGGAAGATCTCCAAATTTCTCAAAACCGATATCGGCCCC 1758

QY 341 GlySerLLeaAlaGluAspValLeuHisAspMetGlyValLLeaAlaMetThrSerSer 360
 DB 1759 GGCTCTATCGCGCTGAAGATGTGCTCCATGATATCGGTGATCGGATGACAAAGTTCC 1818

QY 361 AspSerGlnAlaMetGlyArgAlaGlyGluValLLeaProAspThrTPGlnThrAlaAsp 380
 DB 1819 GATTTCGAAGCAATGGGCGCGCTGGGGAAGTATTTCTTAAGACTTGCGCAAACTGCAAC 1878

QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400


```

QY 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr 240
Db 1352 GAAGACTGGGGCACAACACCAAGTCGATCGATCTGAGGAGTACGAGATGAATAC 1411
QY 241 AspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260
Db 1412 GATGTGCAAGTTTGTATTCACACCGATTCGGTCAATGAGGAGTTATGTATGACACC 1471
QY 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGly 280
Db 1472 CTAATGTGAGTGAACGGGGCGCGCATCTCTACCATTTAGAGGAGCGGGCGGAGA 1531
QY 281 HisSerProAspValIleThrMetAlaGlyLeuAsnIleLeuProSerSerThrThr 300
Db 1532 CACTCACCTGATGTTATCATCATGCGACGACGCTCAATATTCACCTCTCCACGCC 1591
QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
Db 1592 CCCACTATTCCTTATCATTAATACGTTTCAGAACCTTAGACATCTCATGACCTGC 1651
QY 321 HisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340
Db 1652 CACCACCTAGACAAACGATCGCGAGGATCTCCAGTTTCCAAAGCCGATCCGCC 1711
QY 341 GlySerIleAlaAlaGluAspValLeuHisAspMetCysValIleAlaMetThrSerSer 360
Db 1712 GGCTCTATGCGCTGGAAGATGTCTCATATATGCGGATGATCGCATACAAAGCTCG 1771
QY 361 AspSerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAsp 380
Db 1772 GATTGCAAGCAATGGGGCGCGCTGGGGAATGATCTTAAACTTGCCAAATGCGAC 1831
QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspPheArg 400
Db 1832 AAGAAATAAAAAGATTGTGTAACCTCTGAAATGATGTCAGATTAAGCAACTTCGC 1891
QY 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420
Db 1892 ATCAAAAGCTATATCTCCAAATACCACTTAATCCCGTTTGACCCCAAGGGTGAAG 1951
QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValAlaTrpAspProAlaPhe 440
Db 1952 TATATCGGCTCTGCGAAGAGGCAAGATCCCACTTGGTGTGTAATCCCTGCTTC 2011
QY 441 PheGlyValLysProLysIleValIleLysGlyLysMetValAlaPheSerGluMetGly 460
Db 2012 TTGTGTATAAAACCAAAATCGTATCAAAAGCGGTATGCGTGTCTCTGAATGGCC 2071
QY 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480
Db 2072 GATTCTAACCGGTGTGTCACACCTCAGCGGTATTAATCCCGAAATTTTGGGAT 2131
QY 481 HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn 500
Db 2132 CACGCGAAGCGGAATTTGACACACGATCACTTTGTTCCAAAGTCGCTTATGAAT 2191
QY 501 GlyValLysGlyLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsn 520
Db 2192 GGCGTGAAGAAACCTAGGCTTAGACGCGAAGGTGCTACCCGTAAGAAAACTGCCGAC 2251
QY 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys 540
Db 2252 ATCACTAAGAAAGCTTCAATTCACAAACAAAGACGGGCTATACACTGTGATCTTAAA 2311
QY 541 ThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValPro 560
Db 2312 ACCGTGAGGCTTTGTATGATGCAAACTCTGCACCTTAACCCGCTCTGAAGTGCCT 2371
QY 561 LeuAlaGlnArgTyrThrPhePhe 568
Db 2372 CTAGCCCAACGCTACACTTCTTC 2395

```

```

RESULT 4
ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseX subunit encoding sequence #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
FT CDS 2..682
FT FT /*tag= a
FT CDS /*product= "urease protein"
FT FT 693..2399
FT FT /*tag= b
FT FT /*product= "urease protein"
XX
PE EPI176192-A2.
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kueters JG, Catcoli G;
XX
DR WPI; 2002-124384/17.
DR P-PSDB; ADJ58247, ADJ58248.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 10; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseX subunit of the invention.
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 2.97e-259 Length: 2407
Score: 2953.00 Matches: 558
Percent Similarity: 99.12% Conservative: 5
Best Local Similarity: 98.24% Mismatches: 5
Query Match: 98.47% Indels: 0
DB: 7 Gaps: 0
US-09-904-994b-3 (1-568) x ADJ58246 (1-2407)
QY 1 MetLysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
Db 693 ATGAATATGAAAAACAAGATATGTAAATACACGACCTACCAAGCGACAAAGTG 752
QY 21 ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGlu 40
Db 753 CGCTTAGAGATATCCGATCTTTGGCGAAGTGAACATATACCTATATGCGGAA 812
QY 41 GluLeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro 60

```

```

Db      813 GAGCTCAATTGGCGCGGTAAACTATCCGTGAGGCGATGGGTACAGCAATAGTCCA 872
Qy      61 AapGluAerThrLeuAapLeuVal11ethrAapAlaMetC11e1eAapTyrThrGly11e 80
Db      873 GATGAAACACCTTACATTTAGTCATCACCAACGCATGATTATGACATCACCCGGGAAAT 932
Qy      81 TrrLysAlaAapI1eGly11eLysAapNGlyLys11eH1eGly11eGlyLysAlaGlyAapN 100
Db      933 TACAAAGCCGACATTTGGCATTTAAATATGGCAAAATCATAGCATTTGGCAAGGCGAABAC 992
Qy      101 LysAapMetGlnAapGlyValSerProH1sMetVal1aGlyValGlyThrGluAlaLeu 120
Db      999 AAGGACATCAAGATGGCGTAAGCCCTCATATGTCGTGGGTGTGGGACAGAAAGCATTA 1052
Qy      121 AlaGlyGlyGlyMet11e11ethrAlaGlyGly11aAapSerH1sThrH1sPheLeuSer 140
Db      1053 GCGGGGGAAGATATGATTTATACCCGTGGGGGATGATTCACACACCCACTTCCCTCTCT 1112
Qy      141 ProGlnGlnPheProThrAlaLeuAlaAapNGlyVal1ThrThrMetPheGlyGlyGlyThr 160
Db      1113 CACACACATTTCCCTACCCCTTACCCCATAGCGCTTACAAACATGTTTGCGGTGGACAA 1172
Qy      161 GlyProValAapGlyThrAapAlaThrThr11ethrProGlyLysTrrAapLeuH1sArg 180
Db      1173 GGTCCGTGATGATGGACGAATGCGACTACATCATCTCCGGGCAAAATGGAATTTGACCGC 1232
Qy      181 MetLeuAerAlaAlaGlyGlyLysSerMetAapNValGlyPheLeuGlyLysGlyAapSer 200
Db      1233 ATGTTGCGCGAGCTGACAAAGATATTCATGAATGTGGGCTTTTGGGCAAAAGCAATAC 1292
Qy      201 SerSerLysLysGlnLeuValGlyGlnValGlyAlaGlyAla11eGlyPheLysLeuH1s 220
Db      1293 TCCAGTAAACCACTTTGATGACAAATATGAAGCGGGCGCATTCGGCTTTAAATTTGCA 1352
Qy      221 GlnAapTrrGlyThrThrProSerAla11eAapH1sCysLeuSerVal11aAapGlyLys 240
Db      1353 GAAGACTGGGGCAACATCCCAAGTGCATGCATCATCTGCTGAGCGTACAGATGAATAC 1412
Qy      241 AapValGlnValCysAla11ethrAapThrValAapGlnAlaGlyLysValAapAapThr 260
Db      1413 GATGTGCAAGTTTGTATCCACACCGATACGGTCAATGAAGCAGTTATGTAGTGCACAC 1472
Qy      261 LeuAapAlaMetAapNGlyAtrGAla11eH1sAlaTyrH1s11eGlyGlyAlaGlyGly 280
Db      1473 CTGAATGCATGAACCGGGCGCCCATTCATGCTTACCAATGAGGAGCGGGCGAGGA 1532
Qy      281 H1sSerProAapVal11ethrMetAlaGlyGlyLeuAapN11eLeuProSerSerThrThr 300
Db      1533 CACTCAGCTGATGTTATCACATGGCAGGCGAGCTCAATATTTCTACCTTCACACAC 1592
Qy      301 ProThr11eProTyrThr11eAapThrVal11aGlyH1sLeuAapMetLeuMetThrCys 320
Db      1593 CCCACATATCCCTATATCATTTATACGTTTGCAGAACATTTACATGCTCATGACCTGC 1652
Qy      321 H1sH1sLeuAapLysArg11eArgGlyAapLeuGlnPheSerGlnSerArg11eArgPro 340
Db      1653 CACCACTTGATTAACGATCCGCGAGATTTTCAATTTTCCCAAGCCGCTATCCGCCCC 1712
Qy      341 GlySer11eAlaAlaGlyAapValLeuH1sAapMetGlyVal11eAlaMetThrSerSer 360
Db      1713 GGCTCTATCGCCGCTGAAGATGTCTCATGATATTTGGCGTATTCGCAATGACAACTCG 1772
Qy      361 AapSerGlnAlaMetCylArgAlaGlyGlyVal11eProArgThrTrrGlnThrAlaAap 380
Db      1773 GATTTCGAAGCAATGGGGCGCGCTGGCGAAGTATTCCTCGAATCTTGCGAGAT 1832
Qy      381 LysAapLysLysGlyLysPheGlyLysLeuProGlyAapGlyLysAapAapAapAapAap 400
Db      1833 AAGGATTAATAAAGATTTGGTAACTTCTCGAAGATAGTGCACATTAACGAACTTCCGT 1892
Qy      401 H1eLysArgTyr11eSerLysTyrThr11eAapProAlaLeuThrH1sGlyValSerGln 420

```

```

Db      1893 ATCAACGCTACATCTCCAAATATACATTAACCCCGCTTACACCATAGGGGTAAAGCAG 1952
Qy      421 Trr11eGlySerValGlyGlyLys11eAlaAapLeuVal11aTrrAapProAlaPhe 440
Db      1953 TATATCGCTCTGTGTGAAGAGGGCAAAATCGTGAATTTGGTGTGTGAATCTTCCTTT 2012
Qy      441 PheGlyValLysProLys11eVal11eLysGlyGlyMetVal11aPheSerGlnMetGly 460
Db      2013 TTTGGTGTGAACCTTAAGATTTGATCAAAAGCGGATATGATGTCTTCTCTGAATGGGC 2072
Qy      461 AapSerAapAlaSerValProThrProGlnProVal1TrrYArgGlyMetPheGlyH1s 480
Db      2073 GACTCCAAACGCTCCGTGCTACCTCAGCCGCTTATACCGGAAATGTTTGGGCGAT 2132
Qy      481 H1sGlyLysAlaLysPheAapThrSer11ethrPheValSerLysVal11aTrrGluAap 500
Db      2133 CACGGCAAGCGCAAAATTTGACACCAACATCTTTGTTCCAAAGTGTGCTTAAGAAAT 2192
Qy      501 GlyValLysGlyLysLeuGlyLysGlyLysGlnValLeuProValLysAapN11eAap 520
Db      2193 GCGCTGAAGAAAGAACTAGGCTTACAGCGCAAGATGCTACCGTGAATAAAGCGCGCAAC 2252
Qy      521 H1ethrLysLysAapPheLysPheAapLysThrAlaLys11eThrValAapProLys 540
Db      2253 ATCACTAAGAAAGACTTCAAAATTCACAAACAGCGGCGCATATCATCTGATCTTAA 2312
Qy      541 ThrPheGlyValPheValAapGlyLysLeuCysThrSerLysProThrSerGlnValPro 560
Db      2313 ACCTTGAGGCTTTGTGATGTGAACCTGTGACCTTAAACCCGCTTGAAGTGCT 2372
Qy      561 LeuAlaGlnArgTyrThrPhePhe 568
Db      2373 CTAGCCAGCGCTACACTTCTTC 2396

RESULT 5
ADJ58243
ID ADJ58243 strand; DNA; 2183 BP.
XX
AC ADJ58243;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseX subunit encoding sequence #2.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; de.
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
FT CDS 3..683
FT FT /product= "urease protein"
FT FT /tag= b
FT FT /product= "urease protein"
XX
PN EP1176192-AZ.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Catcoul G;
XX
XX
XX MPI: 2002-124384/17.
XX DR P-PSDB; ADJ58244, ADJ58245.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX the diagnosis of Helicobacter felis infections and in the preparation of

```

PT vaccines.
 XX
 PS Disclosure; SEQ ID NO 7; 76pp; English.
 XX
 CC The present invention relates to a novel *Helicobacter felis* urease X and
 CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
 CC used in the manufacture of vaccines against *Helicobacter felis* infections
 CC and in diagnostic tests to detect antibodies against *Helicobacter felis*.
 CC *Helicobacter felis* is difficult to grow so it is more convenient to use
 CC the expression products of the genes encoding the urease X and Y subunits
 CC in the manufacture of vaccines. The present sequence represents an
 CC ureaseX subunit of the invention.
 XX
 SQ Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.26e-226 Length: 2183
 Score: 2588.00 Matches: 489
 Percent Similarity: 99.59% Conservative: 1
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 86.30% Indels: 0
 DB: Gaps: 0
 US-09-904-994b-3 (1-568) x ADJ58243 (1-2183)
 QY 1 MetLysMetLysLysGlnGlnTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
 Db 694 ATGAAATGAAAAAACAAGATATGTAAACCTACGACCCACACAGCGGATAAAGTG 753
 QY 21 ArgLeuGlyAspThrAspLeuThrPAlaGluValGluHisAspTyrThrThrTyrGlyGlu 40
 Db 754 CGCTTAGAGATACCGATCTTGGCGCAAGTAAACATGACTATACCACTTATGGCGAA 813
 QY 41 GluLeuLysPheGlyAlaGlyLysThrIleArgGlnGlyMetGlyGlnSerAsnSerPro 60
 Db 814 GAGTCAAAATTTGGCGGGTAAACATATCCGTAGGGTATGGGTACAGAAATAGCCCA 873
 QY 61 AspGlnAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
 Db 874 GATGAAACACCTTAGATTTAGTATGATACCAACCGCATGATTAATCGACTACCGGAGAT 933
 QY 81 TyrLeuAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100
 Db 934 TATTAAGCCGACATGGTATTTAAATAATGCAAAATCCATGATATTGGCAAGCGGGAGAC 993
 QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeu 120
 Db 994 AAAGACATGCAAGATGCGTAAAGCCCTCATATGGTCGTGGGTGGGCAAGAGCACTA 1053
 QY 121 AlaGlyGlnGlyMetIleIleThrAlaGlyGlyLysAspSerHisThrHisPheLeuSer 140
 Db 1054 GCAAGGGAAGATGATTAATTAACGTCGGGGGATTCATTCGACACCCACTTCTCTCT 1113
 QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyValThr 160
 Db 1114 CCCCAACAAATTCCTTACCGCTTACGCAATGGTITTAACCACTGTTTGAAGTGGACA 1173
 QY 161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArg 180
 Db 1174 GGTCCGGTAGATGCGAAGATGCGACACCATCACTCCGGGCAAAATGAACTTCACCGC 1233
 QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
 Db 1234 ATGTGGCGCAGCTGAAGAGTATTTATGAAATGTAGCTTTTGGCAAGGCAAAATAGT 1293
 QY 201 SerSerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHis 220
 Db 1294 TCTTAGCAAAAACAACCTTGTAGAACAGTAAAGCGGCGCGATGGCTTTAAATTCAT 1353
 QY 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyr 240
 Db 1354 GAAGACTGGGGCAACACCAAGTGCATGCATCATGCTTGAAGCTGGCGAGATGAAATAC 1413

QY 241 AspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260
 Db 1414 GATGTGAAGTTGTATTCACACCGATACGCTCAATAGGAGGAGGTATGTGGATGACACC 1473
 QY 261 LeuAsnAlaMetAsnGlyValGlyAlaIleHisAlaTyrHisIleGlnGlyAlaGlyVal 280
 Db 1474 CTAAATCATGAACCGGCGCGCATCCATCCCTACCACTTAGAGAGCGGCGAGAGA 1533
 QY 281 HisSerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThr 300
 Db 1534 CACTCACTATGATTATCACCATGAGCGAGCGAGCTCAATATTATTAACCTCTCCACACC 1593
 QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
 Db 1594 CCCACTATTCCTATACCATTAATACGTTGCAAGACCTTAGCATGCTCATACCTGCG 1653
 QY 321 HisHisLeuAspLysValArgGlyLeuGlyAspLeuGlnPheSerGlnSerArgIleArgPro 340
 Db 1654 CACCACTTAGATTAAGCCCATCCGAGAGATTTACAAATTTCTCAAAACCGTATCCGCC 1713
 QY 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360
 Db 1714 GATCTATTCGCGCTGAGAGATGTCTCATGATATATGGCGGATCGGATGACTAGCTCC 1773
 QY 361 AspSerGlnAlaMetGlyValGlyAlaGlyValIleProArgThrTrpGlnThrAlaAsp 380
 Db 1774 GATTGCAACCAAGGGGCGCGCTGGGGAAGTATCTTGAAACTTGCGCAATGCGAGAT 1833
 QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAspAspAspPheArg 400
 Db 1834 AAGATTAATAAAGATTTGGTAACTTCTCGAAGATGTGCAGATTAACGAACTTCCGC 1893
 QY 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420
 Db 1894 ATCAAAAGCTATATCTCAATATACACCATTAATCCCTTGAACCCATGCGGTAGCGAG 1953
 QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhe 440
 Db 1954 TATATCGGCTCTGTGAAAGGCGCAAGATGCGCACTTGTGTGGTGAATCTGCGCTTT 2013
 QY 441 PheGlyValLysProLysIleValIleLysGlyGlyMetValIlePheSerGluMetGly 460
 Db 2014 TTTCGCTGAACCTTAAGATTTGTATTAAGTGTGCATGTGTCTCTGAAATGGGC 2073
 QY 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480
 Db 2074 GATCTAAACCGTCGCGCCAGCCCTCAGCCGCTTATTAACGCGAAATGTTTGGGCAC 2133
 QY 481 HisGlyLysAlaLysPheAspThrSerIleThrPhe 492
 Db 2134 CACGCAAGGCGAAATTTGACACCAACATCATCTTT 2169
 RESULT 6
 ID ADQ37850 standard; DNA; 1710 BP.
 XX
 AC ADQ37850;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE H. bizozeronii ureb DNA.
 XX
 XX Urease; urease gene cluster; urease structural gene;
 KW urease accessory gene; ureb; *Helicobacter bizozeronii* infection;
 KW antibacterial; gene; ds.
 XX
 OS *Helicobacter bizozeronii*.
 XX
 FH Key Location/Qualifiers
 FT 1..1710
 FT /*tag= a
 FT /product= "H. bizozeronii ureb"
 XX

QY 462 SerenA1aseValProThrProGlnProValTyrTyrArgGluMetPheGlyHisHis 481
DB 1387 GCCATGCTCTATCCCACTCCCAACCCGTAATTACCCGAAATTTGGCCACCAT 1446
QY 482 G1YLYSA1ALysPheAspThrSerIleThrPheValSerIysValAlaTyrGluAsnGly 501
DB 1447 GGTAAAGCAAAATTGACACCAATATCACTTTGATCCCAAGTGGCTTAGACAAGCC 1506
QY 502 VallyseGluValLeuGlyLeuGluValGlnValLeuProValIysAsnTyrAsnIle 521
DB 1507 ATTTAAAGAGAGTGGCTTCACAAAGAGTGGTTTGCAGTTAAATACTGCCGAACATC 1566
QY 522 ThrIysLYSA1PheLYSA1PheAsnAspLYSA1PheValIleThrValAspProLYSA1 541
DB 1567 ACCAAAGAAAGCTTCAATTAACGATGTTACCGACACATCGAAGTCAATCTGAAGCC 1626
QY 542 PheGluValPheValAspGlyLeuGlyLeuGlySerIysProThrSerGlnValProLeu 561
DB 1627 TACAAAGTTAAAGTGATGCGAAAGAGTTACTTCCAAAGCAGCGAATAATCAGCCTA 1686
QY 562 AlaGlnArgTyrThrPhePhe 568
DB 1687 GCACAACTCTACAACTTGTTC 1707
RESULT 7
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
AC ADQ37847;
XX
DT 07-OCT-2004 (first entry)
XX
DE H. bizozeroni urease gene cluster, ureABIFGH.
XX
KM Urease; urease gene cluster; urease structural gene;
KM urease accessory gene; ureABIFGH; Helicobacter bizozeroni infection;
KM antibacterial; gene; ds.
XX
OS Helicobacter bizozeroni.
XX
PN US2004142343-A1.
XX
PD 22-JUL-2004.
XX
PF 12-AUG-2003; 2003US-00639273.
XX
PR 16-AUG-2002; 2002US-0404337P.
XX
PA (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHU/) ZHU J.
XX
PI Chang Y, Simpson KW, Zhu J;
XX
DR WPI: 2004-533502/51.
DR GENBANK; AF330621.
XX
PT Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizozeroni ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H. bizozeroni.
XX
PS Claim 2; SEQ ID NO 1; 40bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule conferring on
XX Helicobacter bizozeroni an ability to produce urease, where the nucleic
XX acid molecule is a urease gene cluster comprising at least one urease
XX structural gene and at least one urease accessory gene. The nucleic acid
XX molecule is chosen from ureA, ureB, ureC, ureD, ureE, ureF, ureG, ureH and ureI. The
XX invention also relates to an isolated protein encoded by the nucleic
XX acid, a vaccine for preventing onset of disease in mammals infected by H.
XX bizozeroni comprising a nucleic acid and a carrier, and an isolated
XX antibody or its binding portion raised against the nucleic acid. The
XX nucleic acid, proteins and antibodies are useful for vaccinating mammals

CC against onset of disease caused by infection of H. bizozeroni, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizozeroni in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridization assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizozeroni is
CC present in the sample. This sequence represents the H. bizozeroni
CC urease gene cluster, ureABIFGH.
XX
SQ Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,56e-201 Length: 8407
Score: 2326.00 Matches: 420
Percent Similarity: 87.30% Conservative: 75
Best Local Similarity: 74.07% Mismatches: 72
Query Match: 77.56% Indels: 0
DB: 12 Gaps: 0
US-09-904-994B-3 (1-568) x ADQ37847 (1-8407)
QY 2 LysMetLYSA1GlyGluTyrValAsnThrTyrGlyProThrLYSA1AspLYSA1Arg 21
DB 3068 AAAATCTCGAAGAAATATGTTTCTATGATGAGCCCACTAGCGCGATTAAGTAGA 3127
QY 22 LeuGlyAspThrAspLeuTPRAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41
DB 3128 TTGGCGATACCGCATCGATCTTAGAAGTGCAGATGACATGACCATTTAGCGAAGAA 3187
QY 42 LeuLYSA1PheGlyValGlyLeuThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
DB 3188 ATTTAGTTTGGTGGCGGTAAACCATTCGCGATGGATGGACAAACCAAGCCAGC 3247
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81
DB 3248 AGCCACGAATCCATCTTGCTGCTACTAGCCCTGATCGGATTTACCCGGATTTAT 3307
QY 82 LysAlaAspIleGlyIleLYSA1GlyLeuGlyIleHisGlyIleGlyValAlaGlyAsnLYSA 101
DB 3308 AAAGCCGATATGGCATTAATAAATGCGAAATCCATGCGATGGCAAGCAGCAATAAA 3367
QY 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeuAla 121
DB 3368 GACATGCAAGATGGCGTTTGCACCAATCTTGGCTGGCCCTGACTGAGGCTTTGACC 3427
QY 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141
DB 3428 GCTGAAGGCTGATTTTACAGCTGGTGGATTCACACCAATCCACTTATTCTCC 3487
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyIleThrGly 161
DB 3488 CAACAAATCCCAACAGCATTTGGCAGCGGATGACAAACATGATGGTGGGGAACAGGT 3547
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyIleTyrPheAsnIleArgMet 181
DB 3548 CCAAGCTATGGAGACTTAACGACATCAATCCATCCGGGGGCGTGAACCTTAACCATG 3607
QY 182 LeuArgAlaIleGluGlyTyrSerMetAsnValGlyPheLeuGlyLYSA1AsnSerSer 201
DB 3608 CTCGGTCCCTGAAAGAAATATGCCATTAATCTGGGCTTATTTGGGTAAAGGAATGTCT 3667
QY 202 SerLYSA1GlnLeuValGluGlnValAlaGlyValAlaIleGlyPheLYSA1Gly 221
DB 3668 TATGAACCCCTCCCTGGTGCATCAACTGAGCTGAGACCAATGGCTTAATAATCAGAA 3727
QY 222 AspThrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241
DB 3728 GACTGGGGTGAACACCTGAGCATCTACATTTGCTTGAATGGTGAACAATATAGAT 3787
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValIAspAspThrLeu 261
DB 3788 GTGGAATGGCTATCCACACGATACCTTGAATGAAGCGGGCTGTGTGGAAGACATTTG 3847

```

QY 262 AaAlaMeAsnGlyArgAlaIleHisAlaTyRHisIleGluGlyValGluGlyHis 281
DB 3848 CAAGCCATGCTGCGACCTATCCACACTTTCACACTGAGTGTGGCGGCGAC 3907
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrPro 301
DB 3908 GCTCCGGATGTATTAAGATGTCTGGCGAATTAAATCATCTCCACCTTCAACCAACCC 3967
QY 302 ThrIleProTyThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
DB 3968 ACCATTCCCTTTCACCCGTGAATACGAAGCCGAACACATGCAATGTGTGATGTGTCCAC 4027
QY 322 HisLeuAspLeuArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 4028 CACTTGATTAATAACATCAAGAGAGATGTCCAGTTTGTGATTTCTAGATTCGCCCCCAA 4087
QY 342 SerIleAlaAlaGluAspValIleuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 4088 ACCATGCGCGTGAAGACAACACTCCAGATATGGGATTTTCTATCAACAGCTCTGAC 4147
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381
DB 4148 TCCCAAGCATGGCGGTGAGCGAGTCATCACCCGCACTTGGCAACAGCGGACAA 4207
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
DB 4208 AACCAAAAAGATTGGTGTGCTGCTTGAGAAAAGGCGATATACATTCCTCGGCAT 4267
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaIleuThrHisGlyValSerGluTyr 421
DB 4268 AAGCGCTATCATTTCCAAATATACCATCAACCCCGTATATGACACGGCATTTCTGAATAT 4327
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhePhe 441
DB 4328 GTCCGCTCTGTGAAGATGGCAAAATTCGCGATTTGTGCTTGGAGTCTCGGTCTTT 4387
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
DB 4388 GGCATTAAACCCAAACATGATCATCAAGCGGATTCATGCGACTTTCTCAATGGCGAT 4447
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 481
DB 4448 GCCAATGCTCTATCCCACTCCCAACCCGCTATTAACGCGAATGTTGGCCACAT 4507
QY 482 GlyLysAlaLysPheSerThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501
DB 4508 GGTAAAGCCAAATTGACACCAATATCCTTTGTATCCCAAGTGGCTTATGCAACGCGC 4567
QY 502 ValLysGluLysLeuGlyLeuArgGlnValIleuProValLysAsnCysArgAsnIle 521
DB 4568 ATTAAAGAAAGATGGGCTTGCAAGAGTGCTTTGCCAGTAAACACGCCCAACATC 4627
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
DB 4628 ACCAAATAAGACTCTCAATTCACATGATGATCGCACATCATGAAGTCAATCTGTAACC 4687
QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
DB 4688 TACAAAGTTAAAGTGAAGCAAGGATTCTTCCAAAGCAGCGGATTAATCAACGCTA 4747
QY 562 AlaGlnArgTyrThrPhePhe 588
DB 4748 GCACAACTCTACAACTTGTTTC 4768

```

RESULT 8
AAQ04329
ID AAQ04329 standard; DNA; 1710 BP.

XX AAQ04329;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
1

```

DT 21-SEP-1990 (first entry)
XX
XX Probe for sequence encoding part of protein with urease activity.
DE
XX
XX Urease; probe; ss.
KW
XX
XX Helicobacter pylori.
OS
XX
XX WO9004030-A.
PN
XX
XX 19-APR-1990.
PD
XX
XX 06-OCT-1988; 88FR-00013135.
PF
XX
XX 06-OCT-1988; 88FR-00013135.
PR
XX
XX (INSP ) INST PASTEUR.
PA
XX (INRM ) INSERM INST NAT SANTE & RECH MED.
PI
XX
XX Labigne A;
XX
XX MPI, 1990-147844/19.
DR
XX F-PSDB; AAR04580.
XX
XX New nucleotide sequences encoding Campylobacter pylori-ureasents - and
PT derived vectors, transformants, protein, antibodies and probes, useful in
PI diagnosis, treatment and prevention of infections.
XX
XX Claim 7, Page 32; 47p; French.
XX
XX
XX Detection probes consisting of all/part of the sequence can be used to
CC detect C. pylori eg in stomach biopsy samples. The DNA can also be used
CC to express proteins for the prodn. of antibodies and to prepare vaccines.
CC See also AAQ04309, AAQ04311, AAQ04328-30 and AAR04581. (Updated on 25-MAR
CC -2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX SQ Sequence 1710 BP; 537 A; 365 C; 387 G; 421 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,366-199 Length: 1710
Score: 2297.00 Matches: 416
Percent Similarity: 86.77% Conservative: 76
Best Local Similarity: 73.37% Mismatches: 75
Query Match: 76.59% Indels: 0
DB: Gaps: 0

US-09-904-994B-3 (1-568) x AAQ04329 (1-1710)
QY 2 LysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
DB 7 AAGATTACGACAAAGAAATATGTTCTATGTAATGTCTCTACACAGCGGATTAATGAGA 66
QY 22 LeuGlyAspThrAspLeuTyrAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41
DB 67 TTGGCGCATACAGACTTGATGCTGAATAGAAACATGACTACACCATTTATGCGAAGAG 126
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
DB 127 CTTAATATTCGGTGGCGGTAAACCTTAAGAGAGCAAGCAATCTTAACAACTTATGC 186
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81
DB 187 AAAGAAAGACTTGATTTATATTCATCAACGCTTTAATCGGATTAACCGGATTTAT 246
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
DB 247 AAAGCGATATATTGATTAAGATGCAAAATCGCTGCAATGTGAAGCGGTAACAA 306
QY 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeuAla 121
DB 307 GACATGCAAGATGGGCTTAATAAACAATCTTAGCGTAGTCTGCTACTGAAGCCTTAGCC 366

```

QY 122 GYGLUGLWMEITLIEThRALaglyVrLeaSerHsthRHisPheLeuSerPro 141
 DB 367 GGGAGAGTTTGAATGTAACGGCTGGTGTATTAACACATCCATTTATTTACACC 426
 QY 142 GInGInPheProThRALaLeuAlaasnGlyValThrThrmecPheGlyGlyGlyThrGly 161
 DB 427 CAACAAATCCCTAAGCTTTTGCAGAGCGGTTCMAACCAATGATTTGGTGGAGAACCGGT 486
 QY 162 ProValaLePGLyThrAsnAlaThrThrTLeThrProGlyLySTrPAsnLeuHiaArgMet 181
 DB 487 CCGTGCTATGGCACTAATATGCACTACTATCACTCCAGCAGAAATAATTTAAATGATG 546
 QY 182 LeuArgAlaIaIaGUGLyTrSerMetAsnValaGlyPheLeuGlyGlyAsnSerSer 201
 DB 547 CTCAGAGCGGCTGAAGATATTTCTATGAATTTAGGTTTGGCTAAAGGTAAGCGTTCT 606
 QY 202 SerLyLeuGlnLeuValaGUGInValaGlyValaGlyAlaIleGlyPheLyLeuHiaGly 221
 DB 607 AAGATGCGAGCTTAAGCGATCAAAATGAAAGCCGGTGGATTTGGCTTTAAATTCACGAA 666
 QY 222 AspTrpGlyThrThrProSerAlaIleAspHiaScyLeuSerValaIaAspGlyTrAsp 241
 DB 667 GACTGGGGCAACACTCTTCTGCAATCAATCATCGTTAGATGTTGGCGAATAATACGAT 726
 QY 242 ValGInValaCysIleHsthAspThrValaAsnGlyValaGlyTrValaAspAspThrLeu 261
 DB 727 GTGCAAGTCGTAATCCACAGACACTTGAATGAAGCGGTGTGTGAAGAACACTATG 786
 QY 262 AsnAlaMetAsnGlyARGLaIleHiaAlaTyHiaIleGlyValaGlyGlyHia 281
 DB 787 GCGGTATGTGCTGACCCACTATGACACACTTCCACACTGAAGCGCTGGGGGGAGAC 846
 QY 282 SerProAspValIleThrMetAlaGlyGlyLeuAsnIleLeuProSerSerThrThrPro 301
 DB 847 GCTCCTGATATTAATAAGCCGGTGAACACAACATTTCTCCCGCTTCCACTAACCCCC 906
 QY 302 ThrIleProTyThrThrIleAsnThrValaGlyHiaLeuAspMetLeuMetThrCysHia 321
 DB 907 ACCATCCCTTTCACGTAATACAGAAAGACAGACACTGACATGCTTAATGTGTGCAC 966
 QY 322 HiaLeuAspLyAsArgIleARGLaAspLeuGlnPheSerGlnSerArgIleArgProGly 341
 DB 967 CACTGTGATTAAGCATTTAAAGAGATGTTCACTCCGTGATTCAGAGATCCGCCCTCAA 1026
 QY 342 SerIleAlaIaGlyAspValLeuHiaAspMetGlyValaIleAlaMetThrSerSerAsp 361
 DB 1027 ACCATTGCGGCTGAAGACACTTTCATGACATGCGGATTTCTCAATCAACAGATTGAC 1086
 QY 362 SerGlnAlaMetGlyARGLaGlyGlyValIleProArgThrThrProGlnThrAlaAspLy 381
 DB 1087 TCTCAAGCGATGGCCGCTGGGTGAAGTATCACTGAACATTTGGCAAAACGCTGACAAA 1146
 QY 382 AsnLyLeuAspGlyPheGlyLyLeuProGlyAspGlyLyAspAspAspAspPheArgIle 401
 DB 1147 AACAGAAAGATTTGGCCGCTTGAAGAAAGAAAGAACGCAATACGCACTTACGATC 1206
 QY 402 LyAsArgTyIleSerLyTyThrThrIleAsnProAlaLeuThrHiaGlyValaSerGlyTr 421
 DB 1207 AAACGCTACTGTCTAATATCACCATTAACCCAGCATTCGCTCATGGGATTAAGCAATAT 1266
 QY 422 ILGLSerValaGUGLyLyLeuAlaAspLeuValaIleTrpAsnProAlaPhePhe 441
 DB 1267 GTAGGTTCAAGTAAGAGGCAAGAGTGCAGCTGATGTGATGTGAAGTCCAGCAATTCCTT 1326
 QY 442 GYValLyAspProLyIleValaIleLyLeuGlyGlyMetValaIlePheSerGlnMetGlyAsp 461
 DB 1327 GGGGTGAACCAACATGATCATCAAGGCGGATTCATTTGGTTAAAGCAATAGCGGAT 1386
 QY 462 SerAsnAlaSerValaProThrProGlnProValaTyTrArgGlnMetPheGlyHiaHia 481
 DB 1387 GCGAACCCTTCTATCCCTACCCCAACCGGTTTATTAAGAGAAAGTTGTGCTCATCAT 1446
 QY 482 GYLyValaIaLyAspPheAspThrSerIleThrPheValaSerLyValaIaTyrgIuAsnGly 501

DB 1447 GGTAAAGCTTAATATGCATGCAGAACATCACTTTGTGTCTACAGCGCTTATGACAAAAGC 1506
 QY 502 ValLySGlyuLyLeuGlyLeuGluARGLaValLeuProValaLyAsnCyAsArgAsnIle 521
 DB 1507 ATTAAGAGAAATTAAGACTTGAAGAGCAAGTGTGCGGTAAATAATTCAGAAATATC 1566
 QY 522 ThrLyLyAspPheLyPheAsnAspLySThrAlaLySIIeThrValaAspProLySThr 541
 DB 1567 ACTAAAAAGCATGACATTCACACGACACTGCTCACTTGAAGTCMAATCGGAACCT 1626
 QY 542 PheGlyuValPheValaAspGlyLyLeuCyThrSerLyProThrSerGlnValaProLeu 561
 DB 1627 TACCATGTGTGTAAGAGCAAGTAATCTTTAAACAGCAATAAAGTAAAGCTTG 1686
 QY 562 AlaGlnArgTyThrPhePhe 568
 DB 1687 GCGCACTCTTATGACATTTTC 1707
 RESULT 9
 AC34557
 ID AC34557 standard; DNA, 1710 BP.
 AC34557;
 AC 19-JUN-2003 (first entry)
 DT 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #16214.
 DE Antisense; de; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 KM Helicobacter pylori.
 OS MO200277183-A2.
 PN 03-OCT-2002.
 XX 03-OCT-2002.
 PD 21-MAR-2002; 2002MO-US009107.
 PF 21-MAR-2002; 2002MO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haeselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU30687.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 22427; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 613 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)

RESULT 10
AA14230 standard; DNA; 1815 BP.

AA14230;
31-MAR-1999 (first entry)

H. pylori GHPO 1248 gene.
GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
peptic ulcer disease; ss.
Helicobacter pylori.

Key Location/Qualifiers
CDS 51..1780
FT /*tag= a

MO9843478-A1.
08-OCT-1998.
01-APR-1998; 98MO-US006371.
01-APR-1997; 97US-00833457.
24-JUN-1997; 97US-00881227.
29-JUL-1997; 97US-00902615.
(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
(HUMA-) HUMAN GENOME SCI INC.
Kleanthouse H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
P-PSDB; AAM98511.
WPI; 1998-542293/46.
P-PSDB; AAM98511.

New isolated Helicobacter polymnucleotides - used to develop products for the diagnosis, prevention and treatment of Helicobacter infections and gastrointestinal diseases.
Claim 1; Page 960-962; 2054pp; English.

This sequence represents a polymnucleotide of the invention. It was isolated from Helicobacter pylori and encodes a H.pylori GHPO protein. The polypeptides can be used for preventing or treating Helicobacter infections, and gastroduodenal diseases associated with these infections, including acute, chronic, and atrophic gastritis, and peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used for the production of antibodies. The products can also be used for detection and diagnosis

Sequence 1815 BP; 572 A; 388 C; 403 G; 452 T; 0 U; 0 Other;

Alignment Scores:
Pred. No: 1.47e-199 Length: 1815
Score: 2297.00 Matches: 416
Percent Similarity: 86.77% Conservative: 76
Best Local Similarity: 73.37% Mismatches: 75
Query Match: 76.59% Indels: 0
Gaps: 0

US-09-904-994B-3 (1-568) x AA14230 (1-1815)

QY 2 LysMetLysLeuGlnGluValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
DB 57 AAGATTGACGAGAAAGAAATGTTCTATGATGCGCCCTACTACAGCGATAAAGTAGA 116
QY 22 LeuGlyAspThrAspLeuTTPAlaGluValGluHisAspLysThrThrTyrGlyGluGlu 41
DB 117 TTGGCGCATACAGACTGATCGTGAAGTAGAACATGACTACACATTATGAGGAAGAG 176

QY 42 LeuLysPheGlyValGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
DB 177 CTAAATTCGGTGGCGGTAACCTTGAGAGAGAGCATGACCAATCCAAACACCTTACG 236
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspLysThrGlyIleTyr 81
DB 237 AAGAGAAATGATGATCATATCATCATGACCTTAAATCGATGATTACACCGGATTTAT 296
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
DB 297 AAAGCGATATGATTAAGATGAAGATGCAAAATCCCTGGCATGTTAAAGCGGTAACAA 356
QY 102 AspMetGlnAspGlyValSerProHisMetValGlyValGlyThrGluAlaLeuAla 121
DB 357 GACATGCAAGATGGCGTAAACAAATCTTACGATGCTGCTGCTGCTGCTGCTGCTGCTG 416
QY 122 GlyGluGlyMetIleIleThrAlaGlyIleLeuSerHisIleThrHisPheLeuSerPro 141
DB 417 GGTGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 476
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThrGly 161
DB 477 CAACAAATCCCTACAGCTTTTGGCAAGCGGTGTAAACAATGATGATGATGATGATGATG 536
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysSerProLeuHisArgMet 181
DB 537 CTGCTATGCTGCTATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
QY 182 LeuArgAlaIleGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB 597 CTCAGAGCGGCTGAAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATG 656
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221
DB 657 AACGACCGAGCTTACGCGATCAAAATGAAAGCTGCTGCTGCTGCTTAAATCCAGGAA 716
QY 222 AspThrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241
DB 717 GACTGGGACACACTCTTCTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 776
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261
DB 777 GTGCAATCGCTATCCACACAGACACACTTGAATGAATGATGATGATGATGATGATGATG 836
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGlyHis 281
DB 837 GCACTATGCTCGGACGACACTATGCACTTCCACACTGAAGGTGCTGCGGCGGACAC 896
QY 282 SerProAspValIleThrMetAlaGlyGlnLeuAsnIleLeuProSerSerThrPro 301
DB 897 GCTCTATATTTATTAAGTAGCTGTGAACACAACTTCTCCGCTTCCACTAACCCC 956
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
DB 957 ACTATCCCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1016
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 1017 CACTTGATTAAGCATTAAGAAGATGTTCAAGTTCCTGATTAAGATGATGATGATGATG 1076
QY 342 SerIleAlaIleGluAspValIleHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 1077 ACCATTCGCTGAAGCACTTTCATGATGATGATGATGATGATGATGATGATGATGATG 1136
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrProGlnThrAlaAspLys 381
DB 1137 TCTCAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1196
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspAsp 401
DB 1197 AACCAAAAGAAATTTGGCCCTTGAAGAAAGAAAGGCGCATGAACGCAACTTCAGGATC 1256
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProLeuThrHisGlyValSerGluTyr 421

```

Db      1257 AACCGTACTTGTCTAAATACACCATTAACCCAGCATGCGTCATGATTGGCAGATAT 1316
Qy      422 11eglyserValIGluGluGlyLysIleAlaAspLeuValValThrPheProAlaPhePhe 441
Db      1317 GTAGGTTCTGTAGAAATGGCAAGTGGCTGACTTGATTTGGTACGGTCCCATCTTTT 1376
Qy      442 G1ValIysProLysIleValIleLysG1LysIleMetValValPheSerG1LysG1Asp 461
Db      1377 GCGGTAAACCCCAACATGATCATCAAGCGGCTTCATTCGCTGATCAATGGGTGAC 1436
Qy      462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgG1LysMetPheG1LysHis 481
Db      1437 GCGAAGCGTTTATCCCTACCCACACAGATTATTATACAGAGAAATGTCGCTCATCAT 1496
Qy      482 G1LysValAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrG1LysG1 501
Db      1497 GGTTAAGCCAAATACATGCAACATCACTTTGTGTCTCAAGCGGCTTATGCAAAAGGC 1556
Qy      502 ValLysG1LysLeuG1LysLeuG1LysValValLysProValLysAsnG1LysG1AsnIle 521
Db      1557 ATTAAAGAAATTAAGGCTTGAAGACAAAGTGTTCGCGTAAATAATTCAGAAACATC 1616
Qy      522 ThrLysValAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db      1617 ACTAATAAAGACATGCAATTCACACGACCTACGCTCATTCGATGCAATGCAATCTGAAACT 1676
Qy      542 PheG1ValAlaPheValAlaAspG1LysLeuG1LysLeuG1LysSerLysProThrSerGlnValProLeu 561
Db      1677 TACCATGTGTCTGTGATGGCAAAAGTAACTTTCTTAAACGACCAATTAAGTGAAGCTTG 1736
Qy      562 AlaGlnArgTyrThrPhePhe 568
Db      1737 GCGCAACTCTTACGATTTTC 1757

RESULT 11
AAT44351
ID      AAT44351 standard; cDNA; 2735 BP.
XX
XX      AAT44351;
AC
XX      16-OCT-2003 (revised)
DT      11-FEB-1997 (first entry)
XX
XX      H. pylori urea + ureb gene locus in porV214.
DE
XX      Urease; urea gene; ureb gene; vaccine; ds.
XX
XX      Helicobacter; strain CPM630.
OS      unidentified bacteriophage; T7.
OS      Chimeric.
XX
XX
FH      Key
FT      promoter
FT      1.16
FT      /tag= a
FT      /note= "T7 promoter provides transcription initiation for
FT      the urease genes"
FT      33.43
FT      /tag= b
FT      /function= "operator"
FT      /note= "lac operator provides inducible expression of the
FT      urease genes."
FT      complement(46.67)
FT      /tag= c
FT      /note= "B1 primer"
FT      102.818
FT      /tag= d
FT      /product= "urease A subunit"
FT      822.2531
FT      /tag= e
FT      /product= "urease B subunit"
FT      2546.2569
FT      /tag= f
FT      primer_bind

```

```

FT      /note= "BL2 primer"
FT      terminator
FT      2693.2735
FT      /tag= g
FT      /note= "T7 terminator"
XX
XX      W09633732-A1.
XX
XX      31-OCT-1996.
XX
XX      25-APR-1996; 96MO-US005800.
XX
XX      28-APR-1995; 95US-00431041.
XX      PR      06-DEC-1995; 95US-00568122.
XX
XX      (ORAV-) ORAVAX INC.
XX
XX      Lee CK, Monath TP, Ackerman SK, Thomas WD, Kleantous H;
XX      Weltezin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H,
XX      MPI: 1996-497373/49.
XX      P-PSDB; AAM07193, AAM07194.
XX
XX      Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
XX      urease complex and pref. an antibiotic, anti-secretory agent or bismuth
XX      salt.
XX
XX      Disclosure; Page 68-70; 98pp; English.
XX
XX      A cDNA clone (AAT44351), derived from porV214, includes the urea and ureb
XX      sequences coding for the urease A (AAM07193) and urease B (AAM07194)
XX      subunits of Helicobacter pylori clinical isolate CPM630. To obtain
XX      porV214, a genomic DNA library of CPM630 was screened with anti-
XX      Helicobacter urease antibody. A 17 kb SalI fragment from an isolated
XX      clone was subcloned into pUC18 to give pSCPI. PCR primers (AAT44352-53)
XX      were used to amplify a 2.5 kb fragment from pSCPI, which was inserted
XX      into pET24+ to give porV214. The vector was utilized in the prodn. of
XX      recombinant, enzymatically inactive, multimeric urease in E. coli
XX      transformants for use in vaccines to treat or prevent Helicobacter
XX      infection. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX      Sequence 2735 BP; 875 A; 562 C; 635 G; 663 T; 0 U; 0 Other;
SQ
XX

Alignment Scores:
Pred. No.: 2.54e-199 Length: 2735
Score: 2297.00 Matches: 416
Percent Similarity: 86.77% Conservative: 76
Best Local Similarity: 73.37% Mismatches: 75
Query Match: 76.59% Indels: 0
DB: 2 Gaps: 0

US-09-904-994B-3 (1-568) x AAT44351 (1-2735)
Qy      2 LysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysG1AspLysValArg 21
Db      828 AAGATTACACAGAAAGAAATATGTTTCTATGATGCTCTACACAGCGCATTAAGTGA 887
Qy      22 LeuG1AspThrAspLeuTyrAlaGluValGluHisAspTyrThrTyrG1LysGlu 41
Db      888 TTGGCGATACGACTTGTGCTGCTGAAGTAGACATGACTACACCATTTATGCGCAAG 947
Qy      42 LeuLysPheG1LysAlaG1LysThrIleArgG1LysMetG1LysGlnSerAsnSerProAsp 61
Db      948 CTTAAATTCGGTGGCGGTAAACCTTAAGAGAGGACATGAGCATTTAAACACCTTAC 1007
Qy      62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspTyrThrG1LysIleTyr 81
Db      1008 AAAGAAAGATTGGATTATATATCACTAACCGCTTTAATCGTGATTAACCCGATTAT 1067
Qy      82 LysAlaAspIleG1LysLysAsnG1LysIleHisG1LysG1LysValAlaG1LysAsnLys 101
Db      1068 AAAGCGATATTGGATTATTAACATGCAAAATCGCTGGCATTTGTAAGCGGTAAACAA 1127
Qy      102 AspMetGlnAspG1LysSerProHisMetValValG1LysValG1ThrGluAlaLeuAla 121

```

```

Db 1128 GACATGCAAGATGCGCTTAAAAACAATCTTAGCGTAGAGTCTCGCTACTGAAAGCCTTACCC 1187
Qy 122 GYGLUGLWETIIEIETHRALGLYIILEAPSERHIETHRIHPHELEUSERPRO 141
Db 1188 GGTAAAGGTTGATCGTAAACGGCTGGTGTATTGACACACATCACTTCATTCACTTCAACC 1247
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThrGly 161
Db 1248 CAACAATATCCCTACAGCTTTTGGCAAGCGGTAAACACATGATTTGGTGGTGAACCGGCT 1307
Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHiLeuMet 181
Db 1308 CTCTGTATGSCACTAATGCACTACTACTCCAGGACAGAAATTTAAATGATG 1367
Qy 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
Db 1368 CTCAGACCGGCTGAAGAATATTCTATGAAATTAGGTTCTTGGCTAAAGTAAACGCTTCT 1427
Qy 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHiGly 221
Db 1428 AACGATGCGAGCTTAAGCCGATCAATGAAACCGGTGGATTTGGCTTAAATTCACGAA 1487
Qy 222 AspTrpGlyThrThrProSerAlaIleAspHiLeuSerValAlaAspGlyTyrAsp 241
Db 1488 GACTGGGGGCACTCTCTTGCATCAATCAATGATGTTGGCGGACAAATACGAT 1547
Qy 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261
Db 1548 GTGCAGATCGCTATCCACACAGACACTTGAATGAAGCGGTTGGTGAAGACACTATG 1607
Qy 262 AsnAlaMetLeuGlyValGlyAlaIleHisAlaTyrHisIleGluGlyAlaGlyGlyHis 281
Db 1608 GCTGCTATGTGGACGATGACATGACACTTCCACACTGAAGCGGTGGCGGAGAC 1667
Qy 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301
Db 1668 GCTCTGATATTAATAAGTAGCGGCTGAACAAACATTTCTCCGCTTCCACTAACCCC 1727
Qy 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
Db 1728 ACCATCCCTTACCGTGAATACAGAAAGACAGACATGACATGCTTATGCTGTGCAC 1787
Qy 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 1788 CACTTGATTAAGACATTAAAGAGATGTTCACTGCTGATTCAGAGATCCGCTCA 1847
Qy 342 SerIleAlaIleAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
Db 1848 ACCATTCCGCTGAAGACACTTTGCATGACATGGGATTTTCTCAATCACCAATCTTGAC 1907
Qy 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
Db 1908 TCTCAACGATGGGCGCTGTGGTGAAGTATCACTGAACCTTGCAAAACGCTGACAA 1967
Qy 382 AsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
Db 1968 AACCAAGAAAGATTGGCCGCTTAAAGAAAGAAAGCGGCTAAACGAACTTGAAGATC 2027
Qy 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlyTyr 421
Db 2028 AAACGCTACTGTTTAATACATACCACTTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT 2087
Qy 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhePhe 441
Db 2088 GTAGGTTCAGTAGAAAGTGGCGAAAGTGTGATCTTGATGTGAGATCGACATTCCTT 2147
Qy 442 GluValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
Db 2148 GGCGTGAACCCCAACATGATCATCAAGCGGATTCATGCGTTAAAGCAATAGCGGAT 2207
Qy 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 481

```

```

Db 2208 GCGAAGCGTTCTATCCCTACCCCAACCGGTTTATTATACAGAGAATCTTGCCTCATCAT 2267
Qy 482 GYLYSALALYSPhEAPThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501
Db 2268 GGTAAAGCTTAATACGATGACAAACATCATCTTTGTCTCAAGCGGCTTATGACAAAGGC 2327
Qy 502 ValLysGlyLysLeuGlyLysGluValGlnValLeuProValLysAsnCysArgAsnIle 521
Db 2328 ATTAAAGAAATTAAGACTTGAAGAACAAGTGTGCGGTAAATAAATTGCAAGAAATATTC 2387
Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2388 ACTAAAAAGACATGCAATTCACAGACACTACTCTCATTTGAAGTCAATCTGAAACT 2447
Qy 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
Db 2448 TACCATGTGTTGCGGATGCGAAAGAGTAATCTTAACCAAGCAATTAAGTGAGCTTG 2507
Qy 562 AlAGlnArgTyrThrPhePhe 568
Db 2508 GCGCACTCTTACGATTTTC 2528

RESULT 12
AL60579
ID AL60579 standard; DNA; 1710 BP.
AC AL60579;
XX 03-SEP-2003 (first entry)
DT 03-SEP-2003 (first entry)
XX Helicobacter pylori urease B (ureB) DNA.
DE Helicobacter pylori urease B (ureB) DNA.
XX Recombinant protein; plant pollen; urease B; ureB; ds.
XX Helicobacter pylori.
OS Helicobacter pylori.
XX MO2003044050-AL.
XX 30-MAY-2003.
PD 19-NOV-2002; 2002MO-KR002154.
XX 19-NOV-2001; 2001KR-00071712.
XX (PARK/) PARK H.
PA Park H;
PI Park H;
XX WPI; 2003-457590/43.
XX Producing a recombinant protein using plant pollen, for industrial uses,
PT comprises introducing a target gene into an Agrobacterium vector, and
PT infecting the transformed Agrobacterium into a cultured plant pollen.
XX
XX Example 1; Page 29-30; 34pp; English.
XX The invention relates to a method of producing a recombinant protein
XX using plant pollen. The method involves introducing a target gene into an
XX Agrobacterium vector and infecting the transformed Agrobacterium into a
XX cultured plant pollen. The method is useful for producing recombinant
XX proteins for testing, diagnosis and prevention and for industrial use.
XX The present sequence is Helicobacter pylori urease B (ureB) DNA. This
XX sequence is used in the exemplification of the invention
XX
SQ Sequence 1710 BP; 540 A; 371 C; 383 G; 416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.79e-199 Length: 1710
Score: 2291.00 Matches: 415
Percent Similarity: 86.60% Conservative: 76
Best Local Similarity: 73.19% Mismatches: 76
Query Match: 76.39% Indels: 0
DB: Gaps: 0

```


XX Mckenzie B, Boyle J, Lew A;
 XX WPI; 2003-156789/15.
 XX Raising an immune response in an animal by administering composition
 PT comprising carrier and antigen bound to targeting moiety which binds to
 PT receptor present in circulatory vessels in Gut Associated Lymphoid
 PT tissue.
 XX
 PS Disclosure; Page 28-29; 45pp; English.
 XX
 CC This sequence encodes H. pylori urease B. This sequence was used in the
 CC method of the invention for raising an immune response in an animal. The
 CC method comprises administering to the animal a composition comprising a
 CC carrier and an antigen bound to a targeting moiety which binds to at
 CC least one receptor present in circulatory vessels in Gut Associated
 CC Lymphoid tissue (GALT). The method is useful for raising an immune
 CC response in an animal against antigens from Salmonella, Cholera,
 CC Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut
 CC associated toxins, gut hormones, gut hormone receptors or gut associated
 CC cancers. The method is useful for raising both mucosal and systemic
 CC immunity against any antigen used in the composition
 XX
 SQ Sequence 1717 BP; 536 A; 372 C; 388 G; 421 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No: 1,38e-198 Length: 1717
 Score: 2286.00 Matches: 415
 Percent Similarity: 86.42% Conservative: 75
 Best Local Similarity: 73.19% Mismatches: 77
 Query Match: 76.23% Indels: 0
 DB: Gaps: 0
 US-09-904-994B-3 (1-568) x ABA00817 (1-1717)
 QY 2 LysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
 DB 13 AAGATTGACGAGAAAGAAATATGTTCTATGTATGCGCTTACTACAGGATAAAGTGAGA 72
 QY 22 LeuGlyAspThrAspLeuTyrPAlaGluValGluHisAspTyrThrThrTyrGlyGluGlu 41
 DB 73 TTGGCGCATACAGACTTGATGCGCTGAAGTAAACATGACTACACCATTTATGGCGAAG 132
 QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
 DB 133 CTTAAATTCGGTGGCGGTAAGAACTTAAGAGAGAGCATGAGCAATCTAACACCCCTTAC 192
 QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspTyrThrGlyIleTyr 81
 DB 193 AAAGAGAACTGATGATCTATATCATCACTAACCGTTAATCGGATTTACACCGATTTAT 252
 QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysValGlyAsnLys 101
 DB 253 AAAGCGGATATGTTGATTTAAAGATGCGAAATCCCTGSCATTTGGTAAAGCGGTAACAA 312
 QY 102 AspMetGlnAspGlyValSerProHisMetValGlyValGlyThrGlnAlaLeuAla 121
 DB 313 GACTGTGACAGATGCGGTTAAACAACTTGTAGCGGTCTGCTACTGAGAGCCCTTACGC 372
 QY 122 GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
 DB 373 GGTAAAGGTTGATCGTAACCTGCTGGTGATTTGACACACATCCACTTATTTTCAACC 432
 QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyIleThrGly 161
 DB 433 CAAGAAATCCCTACAGCTTTTGCAGCGGTAAACACATGATTTGGTGGCGGAACCTGCT 492
 QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTyrAsnLeuHisArgMet 181
 DB 493 CTTGCTGATGGCAGCTAACGCGACATTACTACCTCAGGTAGAAAGAAATTTTAAATGGATG 552
 QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201

DB 553 CTGAGCGGCTGAAGATTTCTATGAATTTAGCTTTCTGGCTTAAAGGTAAGCTTCT 612
 QY 202 SerLysLysGlnLeuValGluGlnValAlaGlyAlaIleGlyPheLysLeuHisGlu 221
 DB 613 AACGATGCGAGCTTACCGCATCAATTAAGAACCCGCTGCGATTTGCTTTAAATTCACGAA 672
 QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241
 DB 673 GACTGGGCGACACACTCTTCTGCAATCAATCATCTGTTAGATGTGGCGAATAATACGAT 732
 QY 242 ValGlnValCysIleHisThrAspThrValAlaAsnGluAlaGlyTyrValAspAspThrLeu 261
 DB 733 GTGCAAGCTGCTATCCACAGACACACTTGATTAAGCCGCTGTGTGTAAGACACTATG 792
 QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyIleGlyHis 281
 DB 793 GCTGCTATTTGGTGAACCATATGCACACTTCCACACTGAAGCGCTGGCGGACAC 852
 QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrPro 301
 DB 853 GCTCTGATATTTATTAAGTAGCCGGAACACAACTTCTCCCGCTTCACATACCCC 912
 QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisIleAspMetLeuMetThrCysHis 321
 DB 913 ACCATTCCTTTCACCGGAATACAGAAACAGACACATGACATGTTATGTGTGTCAC 972
 QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
 DB 973 CACTTGATTAAGCACTTTAAAGAAAGATTCAAGTTGCTGATTTCAAGATCCGCTCAA 1032
 QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
 DB 1033 ACCATTGCGGTGAAGACACTTTCATGACATGGGATTTTCTCAATCACCAAGTTCTGAC 1092
 QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381
 DB 1093 TCTCAAGCGATGGGCGCTGGGGTGAAGTTATCACTGAATCTGGCAACAGCTGACAA 1152
 QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
 DB 1153 AACAAAGAAAGATTTGGCGCTTGAAGAAAGAAAGAAAGCGGTAACGCACTTACGATC 1212
 QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGluTyr 421
 DB 1213 AAAGCTACTGTTCTAAATATCACCATTAACCCAGCGATTCCTCATGAGCATGAGATAT 1272
 QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValAlaTyrAsnProAlaPhePhe 441
 DB 1273 GTCGTTCTGTGAAGAGGCGAAAGTGAAGCTGATGTGATTTGTGAAGTCCAGATTCCTT 1332
 QY 442 GlyValLysProLysIleValIleLysGlyIleMetValAlaPheSerGluMetGlyAsp 461
 DB 1333 GCGGTAAACCTTAACATGATCATCAAGATGAGGTGATTCGATTAACCAATGGCGAT 1392
 QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgIleMetPheGlyHisHis 481
 DB 1393 GCGAAACCTTCTAACCTTACCCCTCAACCGGTTTATTAACAGAAAGTTGCTGCTACGCT 1452
 QY 482 GlyValAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAlaAsnGly 501
 DB 1453 GGTAAAGCTTAATTAACATGACATCACTTTGTGTCTCAAGCGGCTTTTGAAGAAAGC 1512
 QY 502 ValLysGluLysLeuGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
 DB 1513 ATTAAAGAAATTAAGACTTTGAAGCAAGAGTGTTCGGTAAATAATTCAGAAATCATC 1572
 QY 522 ThrIleLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
 DB 1573 ACCAAAGAAACATGCAATTCACAGACATACCGCTCACTTGAAGTCAATCCGAAACT 1632
 QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561


```

Db 1155 AACAAAAAGGTTGGCGCTTGAAGAGAAAAAGCGCATACGACACTTCGCGATC 1214
Qy 402 Lysargyllysylserlythylthiaseproalaleuthriaglyalseglutyr 421
Db 1215 AAAGGCTACATCTCTAAATACACCATCAACCCCGCATCGGCGATGGATTTCTGACAT 1274
Qy 422 IleglyservalagluglylylelelaaspleuvalaTTPasproalPhephe 441
Db 1275 GTGGGCTCTGTGGAGAGGGGCAATACGCCGACCTCGCTTGGAGTCCGGCTTTCTT 1334
Qy 442 GlyvalysprolylevalilelysglyglymetvalaPheSerGluMetGlyasp 461
Db 1335 GGCAATTAGCCCAATATGATTATTAAGGCGGATTTATGGCTCTCTCAATGGCGCAT 1394
Qy 462 SerasnalaserValProthrProGlnProValtyrtyrarglumetPheglyhihis 481
Db 1395 GCCAATCGCTATTCCACCCCTCAGCCCGCTTATTAACCTGAATGTTTGAACCAT 1454
Qy 482 GlyvalalalysePheapThrSerileThrpheValserlyseValaIatyrGluasngly 501
Db 1455 GGGAAAAACAATTCGACACCAATATCATCTTCTGTTCCAAAGCGGCTTACAGGCGAG 1514
Qy 502 VallysglylyseuglyleuGluargGlnValleuprovalylyasncyargAsnile 521
Db 1515 ATCAAGAGAGACTAGGCGTATGATCGCGTGTATTTGCCAGTGAAGAACTGTGCAATATC 1574
Qy 522 ThriylalyseapPhelelysePheasnasplyThralalyseIethrValasprolythr 541
Db 1575 ACTAAAAAGGACCTCAATTCACGATGTGACCCACATATTGATGTCACACCTGAAACC 1634
Qy 542 PhegluValpheValasplylyleucythrSerlyseProthrSerGlnValProleu 561
Db 1635 TATAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1694
Qy 562 AlaGlnargtyrThrPhephe 568
Db 1695 GCGCACTTATATATTGTTTC 1715

RESULT 15
AAQ12485 ID AAQ12485 standard; DNA; 2767 BP.
XX AC AAQ12485;
XX DT 27-AUG-2003 (revised)
XX DT 23-SEP-1991 (first entry)
XX DE DNA encoding A and B subunits of H. pylori urease.
XX KM Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; ss.
XX OS Helicobacter pylori.
XX FH Key Location/Qualifiers
XX FT CDS 64..717
XX FT /*tag= a
XX FT /label= subunit A
XX FT 721..2400
XX FT /*tag= b
XX FT /label= subunit B
XX FT /note= "31 kD"
XX PN WO9109049-A.
XX PD 27-JUN-1991.
XX PF 19-DEC-1989; 89GB-00028625.
XX PR 19-DEC-1989; 89GB-00028625.
XX PA (THRE-) 31 RES EXPL LTD.

```

```

XX XX
PI Tabagchal IS, Clayton CL, Wren BW, Kleanthous H;
XX WPI; 1991-208084/28.
DR P-PSDB; AAR12515, AAR13550.
XX.
PT Oligo:nucleotide(s) specific for Helicobacter pylori - used as probes and
PT primers to detect H pylori infection, in diagnosis of gastritis, and
PT duodenal and peptic ulceration.
XX
PS Disclosure; Page 16; 28pp; English.
XX
XX The DNA is a 2.7 kb TagI fragment encoding the A and B subunits of H.
CC pylori (previously C. pylori) urease, i.e. the 66 and 31 kD antigens.
CC From the sequence, probes and primers can be designed for the
CC amplification (by PCR) of the gene, to produce a prods. common to all H.
CC pylori strains so far tested and which do not occur in other ureases and
CC can therefore be used as a specific indication of the presence of H.
CC pylori. H. pylori (previously C. pylori) is strongly implicated in the
CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
CC primers/probes can be used for the detection of H. pylori DNA in gastric
CC mucosa, saliva or faecal samples to provide an early diagnosis of
CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
CC OS field.)
XX
SQ Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,01e-195 Length: 2767
Score: 2250.00 Matches: 412
Percent Similarity: 85.71% Conservative: 74
Best Local Similarity: 72.66% Mismatches: 81
Query Match: 75.03% Indels: 1
DB: 2 Gaps: 0

US-09-904-994B-3 (1-568) x AAQ12485 (1-2767)
Qy 2 LysmetlylysginglylythryvalaenthThrtyrGlyprothrlysglyasplyvalarg 21
Db 790 AAGATTGACAGAAAGATATGCTTATGATGAGCCCTACTACAGCGATTAAGTGAAGA 849
Qy 22 LeuGlyaspThrAspleuTPRAlaGluValaGluHisaspTyrrThrtyrGlyGluGlu 41
Db 850 TTGGGCGATACAGACTTGATCGCTGAGTAGACATACATACATCACTTATGCGAAGAG 909
Qy 42 LeuLysPheGlyalaglylyethrileargGluGlymetGlyGlnSerAsnSerProasp 61
Db 910 CTTAAATTCCGGCGCGGTAAACCTTAAGAGAAAGCATGAGCCATCTAACACCTTAGC 969
Qy 62 GluAsnThrleuaspLeuValileThrAsnAlaMetileleaspyrThrGlyiletyr 81
Db 970 AAAGAAAGACTGATCTCAATCATCATCACTTAACGCTTAAATCGATGATTAACCGGATTTAT 1029
Qy 82 LysAlaaspylleGlyilelysaenglylyseIlehisGlylleGlylybAlaGlyAsnlyse 101
Db 1030 AAAGCGATATTGGTATTAAGATGCGAAATGCGTGGCATTTGGTAAAGCGGATTAACAA 1089
Qy 102 AspMetGlnaspGlyValSerProHismetValaGlyValaGlyThrGluAlaLeuAla 121
Db 1090 GACACGCAAGATGGCGTTAAAAACAATCTTAGCGTGGCTCTGCTACTGAAGCCTTAGCC 1149
Qy 122 GlyGluGlymetilelethrAlaGlyGlylyleaspsenrithrHispheluseSerPro 141
Db 1150 GGTGAAGTTGATGTACTGCTGCTGATGATGACACACATCCACTTATCTCCCTCC 1209
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnglyValaThrThrMetPheGlyGlylyThGly 161
Db 1210 CAACAATCCCTACAGCTTTTGCAGGCGGTGAACAACATATGATGGTGGCGAAGTGGC 1269
Qy 162 ProValaspGlyThrAsnAlaThrThrIleThrProGlylyserPheAsnLeuHisArgMet 181
Db 1270 CTTGCTGATGCACTAACGACCAACCATATCATCTCAGGTAGAGAAATTTAAATTCATG 1329

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 20:43:25 ; Search time 5422.47 Seconds

(without alignments)
5075.652 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999
Sequence: 1 MKMKKGEYNTVPTKGDV.....KLCTSKPTSGVPLAQRVTF 568

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09904994/runat.11022005.145647.25534/app.query.fasta_1.1102
-DB=GenEmbl -OPMT=faetaap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=extc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09904994.OCGN.1.1_3980.0/runat.11022005.145647.25534 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pac.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2999	100.0	2883	6	BD185302 Helicobac
2	2999	100.0	2883	6	AX356683 Sequence
3	2973	99.1	2452	6	BD185306 Helicobac
4	2973	99.1	2452	6	AX356695 Sequence

5	2960	98.7	2405	6	BD185303	BD185303 Helicobac
6	2960	98.7	2405	6	AX356686	AX356686 Sequence
7	2953	98.5	2407	6	BD185305	BD185305 Helicobac
8	2953	98.5	2407	6	AX356692	AX356692 Sequence
9	2588	86.3	2183	6	BD185304	BD185304 Helicobac
10	2588	86.3	2183	6	AX356689	AX356689 Sequence
11	2326	77.6	8406	1	AF330621	AF330621 Helicobac
12	2301	76.7	1710	1	AY295085	AY295085 Helicobac
13	2238.5	76.6	2664	1	HECUREASE	HECUREASE
14	2297	76.6	1710	6	A07398	A07398 C.jejuni ge
15	2297	76.6	1710	6	A08818	A08818 C.jejuni DN
16	2297	76.6	1815	6	BD092530	BD092530 Identific
17	2297	76.6	2735	6	AR054309	AR054309 Sequence
18	2297	76.6	5100	1	HECUREABCD	HECUREABCD
19	2297	76.6	7088	1	AB032429	AB032429 Helicobac
20	2297	76.6	10903	1	AE000529	AE000529 Helicobac
21	2297	76.6	12037	1	AE001446	AE001446 Helicobac
22	2292	76.4	1704	1	AY714224	AY714224 Helicobac
23	2291	76.4	1710	1	AF352376	AF352376 Helicobac
24	2288	76.3	2430	1	AY227442	AY227442 Helicobac
25	2281	76.1	4824	6	AR349305	AR349305 Sequence
26	2281	76.1	4824	6	AR349306	AR349306 Sequence
27	2251	75.1	1680	1	AF479026	AF479026 Helicobac
28	2250	75.0	2746	1	HPHPUABG	HPHPUABG
29	2250	75.0	2767	6	A31515	A31515 H.pylori ur
30	2244	74.8	2619	1	HFURE	HFURE
31	2244	74.8	2619	6	A41006	A41006 Sequence 1
32	2244	74.8	2619	6	A72509	A72509 Sequence 1
33	2244	74.8	2619	6	A73285	A73285 Sequence 1
34	2244	74.8	2619	6	AR061045	AR061045 Sequence
35	2244	74.8	2619	6	AR161678	AR161678 Sequence
36	2107	70.3	6126	1	AF332656	AF332656 Helicobac
37	2107	70.3	308929	1	AE017145	AE017145 Helicobac
38	2028.5	67.6	6537	1	BACUREA	BACUREA
39	2012.5	66.1	300150	1	AP001507	AP001507 Bacillus
40	1981.5	66.1	10118	1	BSUREOP	BSUREOP
41	1981.5	66.1	200690	1	BSUB0019	BSUB0019 Bacillus su
42	1960.5	65.4	4485	1	AF035751	AF035751 Synechoco
43	1960.5	65.4	14999	1	AF381045	AF381045 Synechoco
44	1927.5	64.3	106209	1	D90903	D90903 Synechocyst
45	1919	64.0	287298	1	AE017275	AE017275 Bacillus

ALIGNMENTS

RESULT 1	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302				
DEFINITION	Helicobacter felis vaccine.				
ACCESSION	BD185302				
VERSION	BD185302.1	GI:31877502			
KEYWORDS	JP 2002355054-A/1.				
SOURCE	Helicobacter felis				
ORGANISM	Helicobacter felis				
REFERENCE	Bacteria: Proteobacteria; Epsilonproteobacteria; Campylobacterales;				
AUTHORS	Helicobacteriaceae; Helicobacter.				
TITLE	1 (Dases 1 to 2883)				
JOURNAL	Kusters,J.G. and Catolli,G.				
COMMENT	Helicobacter felis vaccine				
	Patent: JP 2002355054-A 1 10-DEC-2002;				
	AKZO NOBEL NV				
	OS Helicobacter felis				
	FN JP 2002355054-A/1				
	PD 10-DEC-2002				
	PF 16-JUL-2001 JP 2001214711				
	PR 17-JUL-2000 EP 00202565.8				
	PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATOLLI				
	PC				
	CI2N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC				
	A61K39/23,				
	PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,CI2N1/15,				
	PC CI2N1/19,				

C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/53,	PC
G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, C12R1:01),	PC
PC C12N15/00, C12N5/00, A61K3/02	
CC Helicobacter	Felis vaccine
PH Key	Location/Qualifiers
FT CDS	(206) . (1886)
FT CDS	(897) . (2603) .

```
/organism="Helicobacter felis"  
/mol_type="genomic DNA"  
/db_xref="taxon:214"
```

ORIGIN

Alignment Scores:

Pred. No.:	3,598-190	Length:	2883
Score:	2999.00	Matches:	568
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
JB:	6	Gaps:	0

US-09-904-994B-3 (1-568) x BD185302 (1-2883)

Qy	1	MethylsMetLysLysGInGluTyrValAlaSerThrTyrGlyProThrLysGlyAspLysVal	20
Db	897	ATGAAAAATGAAAAAACAAGAAATATGTAAATCTACGAGACCACAAAGGGGATTTAACTG	956
Qy	21	ArgLeuGlyAspThrAspLeuTPrpAlaGluValGlnHisAspGlyThrThrTyrGlyGlu	40
Db	957	CGCTTAGAGATACCGATCTTTGGCAGAAAGTAGAACATGACTATACCACTTATGGCGAA	1018
Qy	41	GluLeuLysPheGlyValaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro	60
Db	1017	GAACCTAAATTTGGGGCGGGTAAAACTATCCGTGAGGGTATGGGTCAAGCAATGACCT	1078
Qy	61	AspGluSerThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle	80
Db	1077	GATGAAAAACCCCTAGATTGTAGTCATCAACGGAGAAATTTATCGACTACACCCGGGATT	1138
Qy	81	TyrIleValAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn	100
Db	1137	TACAAAGCCGCATTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGCGAGAAC	1198
Qy	101	LysAspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu	120
Db	1197	AAGGACATGCAGAATGGCGTAAAGCCCTCATATGGTCGATGGGATGGGCACAGAAAGCTA	1258
Qy	121	AlaGlyGluGlyMetIleIleThrAlaGlyLysIleAspSerHisThrHisPheLeuSer	140
Db	1257	GCAGGGGAGAGTATGATTTATTAACCGCTGGGGAAACGATTCACACACCACTCTCTTCT	1318
Qy	141	ProGlnGlnPheProThrAlaLeuAlaAsnGlyValTrrThrMetPheGlyGlyGlyThr	160
Db	1317	CCACACACATTCCTTACCGCTCTAGCCCAATGGCTTACACCATTTTGGAGCGGCACA	1378
Qy	161	GlyProValaAspGlyThrAsnAlaThrThrIleThrProGlyLysTPrpAsnLeuHisArg	180
Db	1377	GGTCCTGTAGATGGCAGCAATGCGCATCTATCACTCCGGGCAATGAACTTGACACCGC	1438
Qy	181	MetLeuAlaGlnAlaGluGluTyrSerMetLeuValGlyPheLeuGlyLysGlyAsnSer	200
Db	1437	ATGTTGGCCCGAGCAGAAAGATATTCTATGATGTGGGCTTTTGGGCAAAAGGCATATGCG	1498
Qy	201	SerSerLysLysGlnLeuValaGluGlnValaGluAlaGlyAlaIleGlyPheLysLeuHis	220
Db	1497	TCTAGCAAAAAACAACCTTGTAAGACAAATAGAAAGCGGGCGCATTTGGTTTAAATTGCAT	1558
Qy	221	GluAspTPrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyThr	240

Db	1557	GAAGACTGGGGCAACAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATAC	1616
QY	241	AspValGlnValCysIleHisThrAspThrValAlaGlnValArgIlyrValAspAspThr	260
Db	1617	GATGGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTATGATGACACC	1676
QY	261	LeuAlaMetAlaGlnIleArgAlaIleHisAlaThrHisIleGluGlnValArgIleGly	280
Db	1677	CTAAATGCAAGAACGGGGCGGCATCCATGCTCCATCCACACTTGAGGAGGGGGTGGAGGA	1736
QY	281	HisSerProAspValIleThrMetAlaGlnIleuLeuAsnIleLeuProSerSerThrThr	300
Db	1737	CACTCACCTGATGTATATCCATGACATGGACGGAGCTCAATATTCTACCTCTCCACACACC	1796
QY	301	ProThrIleProIlyrThrIleAsnThrValAlaGlnIleHisAspMetLeuMetThrCys	320
Db	1797	CCCATATATCCCTATACCATTAATATACGGTTGGACAAACACTTAGACATGCTCATGACATGC	1856
QY	321	HisHisIleuAspIlyrValArgIleArgIleuAspLeuGlnPheSerGlnSerArgIleArgPro	340
Db	1857	CACACCTAGCAAAACGATCCGGAGAGATTTCAAATTTCTCAAAACCGGATCCGCCCC	1916
QY	341	GlySerIleAlaAlaGlnAspValLeuHisAspMetGlyValIleAlaMetThrSerSer	360
Db	1917	GGCTTCATCGGGCTGAAAGATGCTCCATGATATGGGTGTGATGATGCGATGACAGCTCG	1976
QY	361	AspSerGlnAlaMetGlyArgAlaGlnValIleProArgThrTrpGlnThrAlaAsp	380
Db	1977	GATTCGCAGCAATCGGGCGTGCAGGGAGGTATTCCTCGCAACTTGGCAGACTGGCGAT	2036
QY	381	LysAlaIlyrValGluPheGlyLysLeuProGlnAspGlyLysAspAlaAspAsnPheArg	400
Db	2037	AAGATATAAAAAGATTTGTGTAAGCTTCCTGAAACATGGCAAAAGTTACGTAATTTCCGC	2096
QY	401	IleLysArgThrIleSerIlyrThrIleAsnProAlaLeuThrHisGlyValSerGlu	420
Db	2097	ATTAAAGCCTACATCTCCAAATATACATATCAACCCGCTTTGACCCACGGCGTACGGAG	2156
QY	421	TyrIleGlySerValGluGlnGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhe	440
Db	2157	TATATCGGCTGTGGAAAGAGGGCAAGATCGCGCACTTGGTGGTGGATCCGCGCTTT	2216
QY	441	PheGlyValIlyrProLysIleValIleLysGlyGlyMetValValPheSerGluMetGly	460
Db	2217	TTTGGCGTAAACCCCAAAATCGTATCAAAAGCGCGTATGGTGCTTCTCGAAATGGCG	2276
QY	461	AspSerAsnAlaSerValProThrProGlnProValIlyrTyrArgGluMetPheGlyHis	480
Db	2277	GATTCATAACGGGTGTGGCCACCTCCCAACGGGTATTAATACCGCAAAAGTTTGGGCAT	2336
QY	481	HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn	500
Db	2337	CACGGCAAGCGAAATTGACACCAAGCATCACTTTGTTCCAAAAGTCGCTAGAAAT	2396
QY	501	GlyValIlyrGluLysLeuGlyLeuGlnArgGlnValLeuProValIlyrAsnCysArgAsn	520
Db	2397	GGCGTGAAGAAAGCTGGGCTTAGACGCCAAGTCTACCGGTCAAAAATCGCGCTAAC	2456
QY	521	IleThrIlyrLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys	540
Db	2457	ATCCACCAAGAAAGCTTCAGATTCAAGACAAACCGCAAAAATCACCGTCGATCCGAAA	2516
QY	541	ThrPheGlnValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValPro	560
Db	2517	ACCTTCAGAGTCTTTGTATGATGGCAAACTCTGACCTCTAAACCACTCGCAGATGCTT	2576
QY	561	LeuAlaGlnArgTyrThrPhePhe	568
Db	2577	CTAGCGCAAGCGCTACACTTCTTC	2600
RESULT 2			
AK356683			
LOCUS			
2883 bp DNA linear PAT 13-FEB-2002			

DEFINITION Sequence 1 from Patent EP1176192.
ACCESSION AX356683
VERSION AX356683.1 GI:18674020
KEYWORDS
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 Kusters, J.G. and Cattoil, G.
AUTHORS Helicobacter felis vaccine
TITLE Patent: EP 1176192-A 1 30-JAN-2002;
JOURNAL Akzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..2883
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
206..886
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD3255.1"
/db_xref="GI:18674021"
/translation="MKLTPKOEKELLYAGENVARRKRAEGIKLNOPEALAVYSAHIM
DEARRKKTAYQLMEBCMHFLKDEWNPQGNVYDQVTEATFPDGRKLTVMNPIDP
DEHFKAGEVFGCDKDEILNAGEVTELEVTNEGPKSLHVGSHFHEANAKLKEDSE
KAYGRKLDPISGNTLRIGAGQTRAVQDILPGSKKVIQMGVLVNNIADERHKHALDK
AKSHGFIK"
897..2603
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="GI:18674022"
/db_xref="GI:18674022"
/translation="MKMKQOEYVNTYGPTRKDKVRLGDTDLMAVEHDYTYGEBLKE
GAKTIREGQSNSPDENTLIDVTAMIIDYGIKADIGIKNGKIHGIGKAGND
MOGVSPHMYVCGTEALAGEKMITTNGDISHTEPLSPFOFPALANGTTMGGCT
GPVDGTAATTITPEKMLHRLRAEYSIMNVGFLGKSSSKQVLEQVBAKAIIGK
LHEDWGTTPSIDHCLSVADVDVQCITHDYNAGYVDDTLNANMGRALHAHIGS
AGGSHSPDVTIMAGELILPSSITPTIPIYINTVAEHLMDLMTCHLDKIRLDLQPS
OSRIRPGSIAEDVLHDMGVYAMTSSDQSMRAGEVYIPRTWQADNKKKEGLPED
GKNDNRIKRYISKYITNPALITHGVSEIGSVBEKGLADLVANPAPFGVKEPLTVIK
GGMVPSREKMDSNASVTPPPVYIREMFGHGAKKFPTSITPFSKVAIENGVKELG
ERQVLPEKNCRNITKDFKFNDAKITVDPKTEFVVDGLCTSKPTSOVPLAQRVT
F"

ORIGIN
Alignment Scores:
Pred. No.: 3.59e-190 Length: 2883
Score: 2999.00 Matches: 568
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-904-994B-3 (1-568) x AX356683 (1-2883)

QY 1 MetLysMetLysLysGlnGluTyValAsnThrTyArgLysProThrLysGlyAspLysVal 20
Db ATGAAATGAAAAACAAGAAATATGTAAATACCGAACCCACCAAGCGGATAAAGTG 956
QY 21 ArgLysGlyAspThrAspLeuTTPAIGluValGluHisAspTyThrThrTyArgLysGlu 40
Db CGCTTAGAGATACCGATCTTTGGGAGAAAGTAGAATGATGACTATACCACTATGCGAA 1016
QY 41 GluLeuLysPheGlyAlaGlyLysThrIleArgLysGlyMetGlyGlnSerAsnSerPro 60
Db 1017 GAACCTTAATTTGGCGGGGAAAACTATCCGTAGAGGTATGGGTGAGAGAAATAGCCCT 1076
QY 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyThrGlyIle 80
Db 1077 GATGAAACACCCCTTAGTTATGTCATCACTAACCGGATGATTATCGACTACACCGGATT 1136

QY 81 TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyValAlaGlyAsn 100
Db 1137 TACAAAGCCGACATTTGGGATTAATAAAGCGAAATTCATGGCATTTGCCAAGCGGAAAC 1196
QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeu 120
Db 1197 AAGGACATGCAAGATGGCGTGAAGCTTCATATGGCTCGTGGGTGGGACAGAAAGCTTA 1256
QY 121 AlaGlyLysGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140
Db 1257 GCGGGGAGAGTATGATTATTAACCGTGGGGGATGCATTACACACCACTTCTTCTTCT 1316
QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr 160
Db 1317 CCACMAAATTCCTTACCGCTTACGCAAAAGGGGTTCACCAATGTTTGGAGCGGACA 1376
QY 161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTTPAsnLeuHisArg 180
Db 1377 GGTCCTTAGATGGCAGAAATGGCACTACTATCACTCCGGGCAATGAACTTGACACCC 1436
QY 181 MetLeuArgAlaAlaGlyLysTySerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
Db 1437 ATGTTGGCGGACAGAGAAAGTATTCTATGAATGTGGCTTTTGGCAAGGCAATAGC 1496
QY 201 SerSerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHis 220
Db 1497 TCTAGCAAAAACAACCTTGTAAGAACAAAGTGAAGCGGCGCGATGTTTAAATTCAT 1556
QY 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr 240
Db 1557 GAAGACTGGGACACACACCAAGTGCATGCATGATCACTGTTGACCGTGGCAGATGATAC 1616
QY 241 AspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyValAspAspThr 260
Db 1617 GATGTCAAGTTGTATCCACACCGATACAGTCAATGACGAGGATTATGATGACACC 1676
QY 261 LeuAsnAlaMetLeuGlyValGlyAlaIleHisAlaTyrHisIleGlyLysAlaGlyGly 280
Db 1677 CTAAATGCAATGAACGGGCGCGCCATCCATGACCTTACCACTTGGAGGCGGGGAGGA 1736
QY 281 HisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr 300
Db 1737 CACTCACCTATGTTATCACCATGCGAGCGAGCTCAATTTCTACCTCTCCACCAACC 1796
QY 301 ProThrIleProTyThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
Db 1797 CCCACTATTCCTTATACCATTAATATAGGTTGCAAGAACCTTAGACATGATGACATGC 1856
QY 321 HisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340
Db 1857 CACCACCTTAACAACACCATCCCGGAGGATTTCAATTTTCTCAAAAGCCCTATCCGCC 1916
QY 341 GlySerIleAlaAlaGlyAspValIleuHisAspMetGlyValIleAlaMetThrSerSer 360
Db 1917 GGCTCTATCGCGGCTGAAGATGCTTCATGATATGGGTGATGCGGATGACAACTCG 1976
QY 361 AspSerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTPGIntThrAlaAsp 380
Db 1977 GATTCCGACAAATGGGGCGGTGCGAGGAGATGTTCTCGAACTTGGCAGATCGCGAT 2036
QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400
Db 2037 AAGGATTAATAAAGAAATTTGGTAAGCTTCCGAAAGATGCGAAAGATTAACGTAATTTCCG 2096
QY 401 IleLysArgTyIleSerLysTyThrIleAspProAlaLeuThrHisArgValSerGlu 420
Db 2097 ATTTAGCGCTTACTTCTCAATATCACTATCAACCCCGCTTTGACCAACGCGGTGAGGAG 2156
QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValIleTPAsnProAlaPhe 440
Db 2157 TATATGGGCTCTGTGGAAGGCGGACAGATGCGGCACTTGTGTGTGATTCCTGCTTT 2216

QY 441 PheGlyValIysPheProIysIleValIleIysGlyGlyMetValIalPheSerGlyMetGly 460
DB 2217 TTTGGCGTAAACCCAAATCGATCAAAAGGGGATGGTCTTCTCTGAAATGGGC 2276
QY 461 AapSerAsnAlaSerValProThrProGlnProValTyrTyrArgGlyMetPheGlyHis 480
DB 2277 GATTCTAAACGCGCTGTGCGCCACTCCCAACCGGTTATTATACCGGAATGTGTGGGCA 2336
QY 481 HisGlyLysAlaIalysPheAspThrSerIleThrPheValSerIysValAlaTyrGluAn 500
DB 2337 CACGGCAAGGCGAAATTTATACACACATCACTTTGTTCCAAAGCTCCCTATGAAAAAT 2396
QY 501 GlyValIysGlyLysLeuGlyLeuGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 520
DB 2397 GCGGTAAAGAAAGAGCTGGGCTTAAAGCCCAAGTTCTACCGGTCAAAAACCTCCCTAAC 2456
QY 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProIys 540
DB 2457 ATCACCAAGAAAGACTTCAAGTTCAACGACAAACGCAAAATCAACCGTCGATCCGAAA 2516
QY 541 ThrPheGlyValIalPheValAspGlyLysLeuGlyThrSerLysProThrSerGlnValPro 560
DB 2517 ACCTTGAGAGCTTTGTAGATGGCAAACTCTGCACCTTAAACCACTCGCAAGTGCGT 2576
QY 561 LeuAlaGlnArgTyrThrPhePhe 568
DB 2577 CTAGCCCAAGCGCTACACTTCTTC 2600
RESULT 3
BD185306 2452 bp DNA linear PAT 17-JUN-2003
LOCUS Helicobacter felis vaccine.
DEFINITION BD185306
ACCESSION BD185306.1 GI:31877506
VERSION JP 2002355054-A/5.
KEYWORDS Helicobacter felis
SOURCE Helicobacter felis
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2452)
AUTHORS Kusterer,J.G. and Cattolli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 5 10-DEC-2002;
AKZO NOBEL, NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/5
PD 10-DEC-2002
PR 16-JUL-2001 JP 2001214711
PF 17-JUL-2000 EP 00202565, 8
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ 53,
PC G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1:01),(C12Q1/68, PC
C12R1:01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (48)..(728)
FT CDS (739)..(2445).
FEATURES
source 1..2452
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"
ORIGIN
1
+

Alignment Scores:
Pred. No.: 1,58e-188 Length: 2452
Score: 2973.00 Matches: 562
Percent Similarity: 99.65% Conservative: 4
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 99.13% Indels: 0
DB: 6 Gaps: 0

US-09-904-994B-3 (1-568) x BD185306 (1-2452)

QY 1 MetLysMetLysLysGlnGlnIuTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
DB 739 ATGAAAAATGAAAAAACAAGAGTATGTAATACCTACCGGACCCACACAGGCGCATTAAGTG 798
QY 21 ArgLeuGlyAspThrAspLeuThrAlaGluValGluHisAspTyrThrThrTyrGlyGlu 40
DB 799 CGCTTAGAGATACCGATCTTTGGCAGAGTAGAACAATGATATATACCACTTATGCGAA 858
QY 41 GluLeuLysPheGlyValAglyLysThrIleArgGlyGlyMetGlyGlnSerAsnSerPro 60
DB 859 GAACCTCAAAATTCGGTCGACGTAAACCTATCCGTAGGGGTATGGTCAAGCAATAGCCCA 918
QY 61 AapGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
DB 919 GATGAAAAACACTTATGATTATGATGATCAACAACGCGAGATATTTGACTACACCGGGAAT 978
QY 81 TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100
DB 979 TACAAAGCCGACATTTGGCATTTAAATGCAAAATCAATGCGATTGGCAAGGACAGAAAC 1038
QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeu 120
DB 1039 AAGCAATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGTGGGCAACAAGACACTA 1098
QY 121 AlaGlyGlyGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSer 140
DB 1099 GCAGGGGAAAGTATATTATTCCTCGCGGGGATGATGATTCACACACCACTTCTCTCT 1158
QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThr 160
DB 1159 CCACAACATTCCTTACCGCTTACGCAATGGCGTTACACAATGTTGGCGGTGACACA 1218
QY 161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysLeuTrpAsnLeuHisArg 180
DB 1219 GCGCCCGTAGATGGACAGATCGACTACACATCCGCGGCAAAATGAACTTGACCGGC 1278
QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
DB 1279 ATGTTGCGGCGCAGAAAGATATTCATGAAATGTGGGCTTTTGGCAAAAGCAATAGC 1338
QY 201 SerSerLysLysGlnLeuValGluGlnValGluAlaGluAlaIleGlyPheLysLeuHis 220
DB 1339 TCTAGTAAAAAACAATTGTACACAAGTAGAAGGGGGCGCATTTGTTAAATTTGCAT 1398
QY 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyr 240
DB 1399 GAAGACTGGGCGCAACACTCAAGTCGATCGATCACTGCTTGAGGCTGAGCATGAAATAC 1458
QY 241 AapValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260
DB 1459 GATGTGCAAGTTGTATACACCCGATACGGGTACATGAGGCGGTATATGATATACACC 1518
QY 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValAlaGlyGly 280
DB 1519 CTAAATGCAATGAAGGGGGCGCATCCATGCTACCACTTGAAGGAGCGGAGTGAAGGA 1578
QY 281 HisSerProAspValIleThrMetAlaGlyLeuAsnIleLeuProSerSerThrThr 300
DB 1579 CACTGACTGAGTATATACCACTGCAAGGCAAGGATATTTCCACCTCCCTCCCAACAC 1638
QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
DB 1639 CTTACTATCCCTATATACATTAATACGTTGCAAGCACTTATGACATGCTTATGACCTGC 1698


```

QY 161 GIYPovaIAspGIyThraSnaIAThrThrIleThrProGIyLysTrpAsnLeuHISaG 180
DB 1219 GGGCCCGTGAAGGCGAAGTCGACTACATCCTCCGGCAAAATGAACTTGACACCC 1278
QY 181 MetLeuArgAlaIaGIuGIuTySerMetAsnValGIyPheLeuGIyLysGIyAsnSer 200
DB 1279 ATGTTGCGCGGACGAGAAAGATTTCTAATGATGTGGCTTTTGGGCAAAAGCAATAC 1338
QY 201 SerSerLysLysGIuLeuValGIuGIuValaGIuValaIleGIyPheLysLeuHIS 220
DB 1339 TCTAGTAAAAAACCACTTATAGAACAGTAGAAGCGGGCGCAATGGCTTTAAATTGCAT 1398
QY 221 GIUAAPTrpGIyThrThrProSerAlaIleAspHisCysLeuSerValaIaAspGIuTy 240
DB 1399 GAAAGCTGGGGCAACACTCCAGTCGATCGATCGCTTGAGCGCTACAGATGAATAC 1458
QY 241 AApValGIuValCysIleHisThrAspThrValAsnGIuValaGIyTyValaAspAspThr 260
DB 1459 GATGTGCAAGTTGTGTACACACCGATACGGTCAATGAAGGAGTTATGTAGACACAC 1518
QY 261 LeuAsnAlaMetAsnGIyArgAlaIleHisAlaTyHisIleGIuGIuValaGIyGIy 280
DB 1519 CTAAATGCAATAGACGGCGCGCATCGCTACACCAATGAGGAGCGGGTGAAGGA 1578
QY 281 HisSerProAspValIleThrMetAlaGIyGIuLeuAsnIleLeuProSerSerThrThr 300
DB 1579 CACTCACCTGATGTTATACCATGGAGCGAAGTAATTTCTACCTCCTCCACCAAC 1638
QY 301 ProThrIleProTyThrIleAsnThrValaIaGIuHisLeuAspMetLeuMetThrCys 320
DB 1639 CCTACTATCCCTATACCACTTAATACGTTTGCAGAACACTTACACTTGTATGACCTGC 1698
QY 321 HisHisLeuAspLysArgIleArgIuAspLeuGIuHisSerGIuSerArgIleArgPro 340
DB 1699 CACCACTAGATTAAGCAATCGCGAGGATCTCCAAATTTCTCAAAAGCGGATCGCCCC 1758
QY 341 GIySerIleAlaIaGIuAspValLeuHisAspMetGIyValIleAlaMetThrSerSer 360
DB 1759 GGCTCTATCGCGCTGAAGATGTGCTCCATGATATCGGTGTATCGCGATGACAGATTCC 1818
QY 361 AAspSerGIuAlaMetGIyArgAlaGIyGIuValIleProArgThrTrpGIuThrAlaAsp 380
DB 1819 GATTCGCAAGCAATGGGGCGCGCTGGGGAAGTATCTTGAACCTTGGCAAACTGCAGAC 1878
QY 381 LysAsnLysLysGIuPheGIyLysLeuProGIuAspGIyLysAspAsnAspAsnPheArg 400
DB 1879 AAGAATAAAAAGAAATTTGGTAAAGTTCTCGAAGATGTGCAATATAGCAACTTCCGC 1938
QY 401 IleLysArgTyrlSerLysTyThrIleAsnProAlaLeuThrHisGIyValSerGIu 420
DB 1939 ATCAAAACGCTATATCTCCAAATACACCATTAATCCCGCTTGAACCATGGCGTAGCGAG 1998
QY 421 TyrlIleGIySerValGIuGIuGIyLysIleAlaAspLeuValaITrAsnProAlaPhe 440
DB 1999 TATATCGCTCTGTGGAAAGGGGCAAGATCGCCGACTTGTGTGGAATCCTGCTTT 2058
QY 441 PheGIyValLysProLysIleValIleLysGIyGIyMetValaIaPheSerGIuMetGIy 460
DB 2059 TTTGGCGTAAACCCAAATTCGATCAAAAGGGGATGTGTGTCTTGAAATGGGCG 2118
QY 461 AAspSerAsnAlaSerValProThrProGIuProValTyTrArgGIuMetPheGIyHis 480
DB 2119 GATTCMAATGGCTGTGGCCACTCCTCAAGCGGTTTATTAACGGGAAATGTTTGGGCAT 2178
QY 481 HisGIyLysAlaLysPheAspThrSerIleThrPheValSerLysValaIaTyrgIuAsn 500
DB 2179 CACGGCAAGCGCAAAATTTGACACCAACATCACTTTGTTTCCAAAGTCCCTATGAAAAAT 2238
QY 501 GIyValLysGIuLysLeuGIyLeuGIuValArgAlaIleuProValLysAsnCysAspAsn 520
DB 2239 GGTTGCAAAAGAAACCTAGGTTTGAAGCGGAAAGGTCTCCCGTGAAAAACGCGCTAAC 2298
QY 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValaAspProLys 540

```

```

DB 2299 ATCAACCAAGAGAACTTCAAGTTCAACGACAAACTGCAAAAAATCACCGTCGATCGAAA 2358
QY 541 ThrPheGIuValPheValaAspGIyLysLeuCysThrSerLysProThrSerGIuValPro 560
DB 2359 ACCTTCGAGGCTTTGTGTGAATGGCAAACTGTGACCTCTAAACCACTCTGAAGTGCTT 2418
QY 561 LeuAlaIaArgTyThrPhePhe 568
DB 2419 CTAGCCCAACGCTACACTTTCTTC 2442

RESULT 5
BD185303 2405 bp DNA linear PAT 17-JUN-2003
LOCUS BD185303
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185303
VERSION BD185303.1 GI:31877503
KEYWORDS JP 2002355054-A/2.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
1 (bases 1 to 2405)
REFERENCE
1 Kusters,J.G. and CatoIi,G.
AUTHORS Helicobacter felis vaccine
TITLE Patent: JP 2002355054-A 2 10-DEC-2002;
JOURNAL AKZO NOBEL NV
COMMENT
OS Helicobacter felis
PN JP 2002355054-A/2
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/
53,
PC G01N33/53,G01N33/566,G01N33/569//C12N9/80,C12R1/01,C12Q1/68, PC
C12R1/01,
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (1)..(681)
FT CDS (692)..(2398)
FEATURES
source location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Alignment Scores: 1,13e-187 Length: 2405
Score: 2960.00 Matches: 560
Pctid: 99.30% Conservative: 4
Percent Similarity: 98.59% Mismatches: 0
Best Local Similarity: 98.70% Indels: 0
Query Match: 6 Gaps: 0
DB: 6

US-09-904-994B-3 (1-568) x BD185303 (1-2405)
QY 1 MetLysMetLysLysGIuGIuTyValAsnThrTyrgIyProThrLysGIyAspLysVal 20
DB 692 ATGAAAAAGAAAAACAAGATATGTAACACTTACGAGACCAACAGGCGATATAAGTG 751
QY 21 ArgLeuGIyAspThrAspLeuTrpAlaGIuValaGIuHisAspTyThrThrTyrgIyGIu 40
DB 752 CGCTTAGAGATACGATCTTTGGGCAAGAAAGTAACAATGACTATACCACTTATGCGGAA 811

```

QY 41 GluLeuLeuSpheGlyAlaGlyLeuThr11LeuArgGluGlyMetGlyGlnSerHisSerPro 60
 Db 812 GAGCTCAAAATTTGGCGGGTAATAAATCCGTGAGGGTATGGGCAAGAAATAGGCCA 871
 QY 61 AspGluSerThrLeuAspLeuVal11LeuHisAlaMetC11LeuAspTyrThrGlyLe 80
 Db 872 GATATAAACAACCTTATGATTTATGATATCCAAACCGGATGATTTATGATCAACCGGGATT 931
 QY 81 TyrLeuAlaAsp11LeuGlyLeuAsnGlyLeuSer11LeuGlyLeuAlaGlyAsn 100
 Db 932 TATATAACCCGCACTTGGTATTAATAATGCAAAATCCATGATTTGGCAAGCGGGGAAAC 991
 QY 101 LeuAspMetGlnAspGlyValSerProHisMetVal11GlyValGlyThrGluAlaLeu 120
 Db 992 AAAACAATGCAAGATGGCGTAAAGCCCTCATATGTCGGGTGGGCAAGAACACTA 1051
 QY 121 AlaGlyGluGlyMet11LeuThrAlaGlyLeuAspSerHisSerHisSerHisSer 140
 Db 1052 GCAGGGGAAGGTATGATTTATTAACCGCTGGGGGATCCATTTGCACACCCACTTCTCT 1111
 QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyVal11ThrThrMetPheGlyGlyThr 160
 Db 1112 CCCCAACAATCCCTACCGCTCTAGCCCAATGATGTTACCAACATGTTTGAAGTGGACA 1171
 QY 161 GlyProVal1AspGlyThrAsnAlaThrThr11LeuProGlyLeuGlyThrAsnLeuHis 180
 Db 1172 GGTCCGGTAGATGGCAGAAATGCAACCACTCATCTCCGGGCAATGGAACTTGACCGC 1231
 QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnVal11GlyPheLeuGlyLeuGlyAsnSer 200
 Db 1232 ATGTGGCGCGAGCTGAAGATATTTCTATGATGTTGGGCTTTTGGCAAGGCAATAGC 1291
 QY 201 SerSerLeuLeuGlnLeuVal11GluGlnVal11GlyAla11GlyPheLeuHis 220
 Db 1292 TCCAGTAAAAACAACCTGATGAAACAAGTAGAAGCGGGCGGATGGCTTTAAATTTGAT 1351
 QY 221 GluAspTrpGlyThrThrProSerAla11LeuAspHisSerLeuSerVal11AspGlyTyr 240
 Db 1352 GAAGACTGGGGCAACCAACCAAGGCGATCGATCATCTTGACCGTGAAGTAAATAC 1411
 QY 241 AspValGlnValCys11LeuSerAspThrVal11AsnGluAlaGlyTyrVal11AspAspThr 260
 Db 1412 GATGTGCAAGTTTGTATCCACACCGATACCGGTCAATGAGGATTTATGATGACACC 1471
 QY 261 LeuAsnAlaMetAsnGlyArgAla11LeuAlaTyrHis11GluGlyAlaGlyGly 280
 Db 1472 CTAAATGCGATGAACGGGGCGGCATCATCTTACCAATGAGGAGCGGGCGAGAGA 1531
 QY 281 HisSerProAspVal11LeuMetAlaGlyGluLeuAsn11LeuProSerSerThrThr 300
 Db 1532 CACTCATCTATGTTATCACCATGCGAGGCGACGCTCATATTTCACTCTCCACCAACC 1591
 QY 301 ProThr11LeuProTyrThr11LeuSerThrVal11AlaGluHisLeuAspMetLeuMetThrCys 320
 Db 1592 CCCACTTTTCCCTATACCATTTATACGTTGCAAAACATTAGACATGCTATGACCTGC 1651
 QY 321 HisHisLeuAspLeuArg11LeuArgGluAspLeuGlnPheSerGlnSerArg11LeuArgPro 340
 Db 1652 CACCACTTAGACAAACCGCATCCGAGAGATCTCCAGTTTCCCAAGCCGATCCGCCCC 1711
 QY 341 GlySer11LeuAlaGluAspVal11LeuHisAspMetGlyVal11LeuAlaMetThrSerSer 360
 Db 1712 GGCTCTATTTCCGCTGAAGATGCTCATATATGCGCGATGCGGATGCAAGGCTCG 1771
 QY 361 AspSerGlnAlaMetGlyArgAlaGlyVal11LeuProArgThrTrpGlnThrAlaAsp 380
 Db 1772 GATTTCGAAACGATGGGGCGGCTGGGGAACTGTTCTTGAACCTTGCAAACTGCGAGC 1831
 QY 381 LeuAsnLeuLeuGluGluGluLeuProGluAspGlyLeuAspAsnAspAspAspArg 400
 Db 1832 AAGAAATAAAAAAGAAATTTGGTAAGCTTCTGAAATGATGATGATTAATGACCACTTCGC 1891

QY 401 11LeuArgTyr11LeuSerTyrThr11LeuAsnProAlaLeuThrHisGlyValSerGlu 420
 Db 1992 ATCAACGCTATATCTCCAAATACACCAATTAATCCGCTTTGACCCATGGGTAGCGAG 1951
 QY 421 Tyr11LeuSerVal11GluGluGlyLeu11LeuAspLeuVal11TPAsnProAlaPhe 440
 Db 1952 TATATCCGCTCTGTGGAAGAGGCAAGATCCGCCACTTGGTGGTGGAAATCTCCTTTC 2011
 QY 441 PheGlyValLeuProVal11LeuVal11LeuGlyGlyMetVal11PheSerGluMetGly 460
 Db 2012 TTTGGTGTAAAAACCAAAATGATGATCAAGGCGGATAGTGGTGGTCTTGAATAGGCG 2071
 QY 461 AspSerHisAlaSerVal11ProThrProGlnProVal11TyrTyrArgGluMetPheGlyHis 480
 Db 2072 GATTCTAACCGGTGTGGCCACACTCAGCGGGTTATTAACCGCAAAATGTTGGGCAT 2131
 QY 481 HisGlyValAlaLeuPheAspThrSer11LeuThrPheValSerLeuValAlaTyrGluAsn 500
 Db 2132 CACGGCAAGGGAATTTGACACACGATCACTTTGTTCCAAAGTCGCTTATGAAAT 2191
 QY 501 GlyVal11GlyLeuGlyLeuGlyLeuGluArgGlnVal11LeuProVal11LeuAsnCysArgAsn 520
 Db 2192 GCGGTAAAAAGAAATAGGCTTAGCGCAAGGCTGTAACCGTGAATAACTGCGCAAC 2251
 QY 521 11LeuThrLeuAspPheLeuPheAsnAspLeuThrAlaLeu11LeuThrVal11AspProLeu 540
 Db 2252 ATCACTAGAAAGCTTCAATTCACAAACAAAGCGGCAATACATGCTGATCTGATAA 2311
 QY 541 ThrPheGluValPheVal11AspGlyLeuGlyCysThrSerLeuProThrSerGluValPro 560
 Db 2312 ACCTTCAGAGTCTTTGTAGATGGCAAACTTGCACTTAAACCCGCTCTGAAGTGCT 2371
 QY 561 LeuAlaGluArgTyrThrPhePhe 568
 Db 2372 CTAGCCCAACGCTACACTTCTTC 2395

RESULT 6
 AX356686
 LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 4 from Patent Epi1176192.
 ACCESSION AX356686
 VERSION AX356686.1 GI:18674023
 KEYWORDS
 SOURCE Helicobacter felis
 ORGANISM Helicobacter felis
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteraceae; Helicobacter.

REFERENCE
 1 Kusters,J.G. and Cattolli,G.
 Helicobacter felis vaccine
 Patent: EP 1176192-A 4 30-JAN-2002;
 Akzo Nobel N.V. (NL)

FEATURES
 source
 1..2405
 Location/Qualifiers
 /organism="Helicobacter felis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:214"
 1..681
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD23257.1"
 /db_xref="GI:18674024"
 /translation="MKLTPEQKFLLYAGAVARKKAGLKLINPEALVAYISAHIM
 DEARRGKTVALEMECMHPLKKDQVYGVGVNVPDLYGVAEPDQTKLVNMPLEP
 DEHPKGEVYFGDKDIEINAGKEVELEVTNPGPSLHNGSHHPEPTKALKFDSE
 KAYGKLDIPSSNTLRIAGQTRKNQILPLGSSKRYIGMNGLVNMLADEHHKKAIDK
 AKSHGRFK"

CDS
 692..2398
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD23258.1"

```

/db xref="GI:18674025"
/translation="MKMKQKQEVVNTYGPPTGDKVRLGPDMLAEVHDYTTYGEELKE
GAGKTRBEGSGNSNPENTLILVITNMIIDYTGIXKADIGIKNGKHGIGKANKD
MODGVSPHNVGVTEALAGEBMTITAGIDSHTEPLSQPITALANKVTTFGGGT
GVDGDTNATITTPGKNLHRMLRAAEYSMNVPFLQKSSKQDVEVEGALGFK
LHEDGTPSADIDHCLSVADYDVOVCHITDYNAGYDDTLNANNGAIAHYHIEG
AGGSGSPVITVAGELNLPSSSTPTPIPTVTAHEHLDMLTCHLDRIRIEDLOFS
OSRIRPGSIAADVDLHDIGVIMTSSDSOGMAGAEVIRPTQTDKNKKEKGPED
GADNPNFRKIKYKTYTNIPALTHGVSEKIDYVWNPAPRGVEXKVIK
GGMVVPSEKGDNSAVPTPQPTTYRMRPFHKGAKPDTSTTVSKVAIENGKYEKGL
ERRVLPVKNCRNITKKDFKFNKTAHIITVDPKTFEYVDGKLCTSKPASEVPLAORT
FP"

```

ORIGIN

Alignment Scores:

Pred. No.:	1,13e-187	Length:	2405
Score:	2960.00	Matches:	560
Percent Similarity:	99.30%	Conservative:	4
Best Local Similarity:	98.59%	Mismatches:	4
Query Match:	98.70%	Indels:	0
DB:	6	Gaps:	0

US-09-904-994b-3 (1-568) x AX356686 (1-2405)

```

QY 1 MetLysMetLysLeuGlnGluValAsnThrTyrGlyProThrLysGlyAspLysVal 20
Db 692 ATGAAAATGAAAAACAAAGATGTAATAACCTTACGACCCACCAAGGATTAAG 751
QY 21 ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGlu 40
Db 752 CCTTAGGAGATACCATCTTGGGCGAGAGTAGAACATGACATATCCACTTAATGCGAA 811
QY 41 GlnLeuLysPheGlnValAlaGlyLysThrTyrLeuGlnGlyMetGlyLysIleAsnSerPro 60
Db 812 GAGCTCAAAATTTGGCGGGTAAACTATCCGAGAGGATGAGGTAGAGCAATACCCA 871
QY 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
Db 872 GATGAAAACACCTTATATTAGATGATCAACAAGCATGATTATGACATACCCGGGATT 931
QY 81 TyrLysAlaAspTyrLeuGlyLysAsnGlyLysIleHisGlyTyrLeuGlyLysAlaGlyAsn 100
Db 932 TATAAAGCCGACATTCGTATTAATAATGCAAAATCCATGATGATGCGAAAGCGGGAAC 991
QY 101 LysAspMetGluAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu 120
Db 992 AAGAATCATCAGATGCGTAAAGCCTCATATGCTGAGGTGGGCAAGAGCATTA 1051
QY 121 AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140
Db 1052 GCAAGGGGAAGGATATGATTATACCGTGGGGGATCGATTCCGACACCCACTTCCCTCT 1111
QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThr 160
Db 1112 CCCCAACAATTCCTCTACCTCCCTACGCAATGGGTGTACAAACATGTTTGGAGGTGCACA 1171
QY 161 GlyProValAspGlyThrAsnAlaThrTrpTyrIleThrProGlyLysThrAsnLeuHisAsp 180
Db 1172 GGTCCGGTGAATGCGCAAGATGCAACCATCATCCGGGCAAAAGAACTTGACCCGC 1231
QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
Db 1232 ATGTTCCGGCGACCTGAAGGATTTCTATGAATGTTGGGCTTTTGGGCAAAAGCAATAC 1291
QY 201 SerSerLysLysGlnLeuValGluGlnValGluAlaGlyValAlaIleGlyPheLysLeuHis 220
Db 1292 TCACATTAATAAAACAACCTCTAGAAAGTAAGAACCGGGCGCGATTGGCTTTAATATGCAT 1351
QY 221 GlnAspTyrGlyTyrThrProSerAlaAlaAspHisCysLeuSerValAlaAspGluTyr 240
Db 1352 GAAGACTGGGGCAACACCAAGTGGATGATCATCGCTTGAGCGGTAGAGATGAATAC 1411
QY 241 AspValGlnValCysIleHisThrAspThrValAsnGluValAlaGlyTyrValAspAspThr 260

```

```

Db 1412 GATGTCGAAGTTTGTATCCACACCGATACGGTCATATGAGGCGATTATAGTATGACACACC 1471
QY 261 LeuAsnAlaMetAsnGlyAlaArgAlaIleHisAlaTyrHisIleGluGlyValAlaGlyLys 280
Db 1472 CTAAATGGATGAATAAGGGGGCGCCCATCATGCTTACCATTTGAAGGAGGGGGCGAGGA 1531
QY 281 HisSerProAspValIleIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThr 300
Db 1532 CACTCACTGATGTTATATCAATGCGACGCGACGCTCATATTTTACCTCCCTCCACAC 1591
QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLysAspMetLeuMetThrCys 320
Db 1592 CCCACTATTTCCCTATACCATTTATACGTTGTGCAAGAACCTTAGATGCTCATGACCTGC 1651
QY 321 HisHisLysLeuAspLysAlaGlyTyrLeuGlnLeuAspLeuGlnPheSerGlnSerArgIleArgPro 340
Db 1652 CACCACTTAGAACAAACGCAATCCGCGAGATCTCCAGTTTTCCAAAGCCGTATCCGCCCC 1711
QY 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360
Db 1712 GGCTCTATTGCGCGTGAAGATGCTCATATATTTGGCGTATGCGCATGACCAAGCTCG 1771
QY 361 AspSerGlnAlaMetGlyAlaArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAsp 380
Db 1772 GATTGCAAGCAATGGGGCGCGCTGGGGAAGTGAATTCCTAGAACTTGGCAACTGCAGAC 1831
QY 381 LysAsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspAsp 400
Db 1832 AAGAAATAAAAAGAAATTTGGTAGCTTCTTGAAAGTGGGAGATATACAACTTCGCC 1891
QY 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420
Db 1892 ATCAACCGTATATCTCCAAATACACATTAATCCCGCTTACCCATGCGCTGAGCGAG 1951
QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhe 440
Db 1952 TATATCGGCTGTGGAAGAGGGCAAGATCGCCGACTTGTGTGTGGAATCTGCTTTC 2011
QY 441 PheGlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGlnMetGly 460
Db 2012 TTTGGTGTAAACCCCAAAATCGTATCAAAAGCGGTATGCTGTCTTCTTAAATGGGC 2071
QY 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480
Db 2072 GATTCTAACGCTCTGTGCGCACACTCAGCCGGTTTATACGGGAATGTTGGGCAT 2131
QY 481 HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn 500
Db 2132 CACGGCAAGCGAAATTTGACACACGATCACTTTTGTTCAAAGTCGCTATGAAAT 2191
QY 501 GlyValLysGlnLysLeuGlyLysLeuArgGlnValIleAspProValLysAsnCysArgAsn 520
Db 2192 GCGGTGAAGAAACCTAGGCTTAGAGCGCAAGGTGTACCCGTAAACCTGCGCAAC 2251
QY 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys 540
Db 2252 ATCATTAAGAAAGACTTCAAAATTCACAAAGAGGGCGCATACACTGTGATCTTAA 2311
QY 541 ThrPheGluValPheValAspGlyLysLysCysThrSerLysProThrSerGlnValPro 560
Db 2312 ACCTTCGAGGTCTTGTGATGAGCAAACTGTGCACTCTAAACCCGCTGTAAGTGCCT 2371
QY 561 LeuAlaGlnArgTyrThrPhePhe 568
Db 2372 CTAGCCCAACGCTACACTTCTTC 2395

```

RESULT 7
 LOCUS BD185305 2407 bp DNA linear PAT 17-JUN-2003
 DEFINITION Helicobacter felis vaccine.
 ACCESSION BD185305
 VERSION BD185305.1 GI:31877505

Db 2133 CACGGCAAGCCGAATTTGACACGACATCATCTTTGTTTCCAAAGTCCCTATGAAAT 2192
Qy 501 GtYvAllySGluLySleuGlyLeuGluValgInValleuProValLySaNCySaTgaen 520
Db 2193 GCGCGGAAAGAAACCTAGCGCTTAGCGCCAGAGTCTCCCGTGAAAAACGCCCAAC 2252
Qy 521 lIeThrLyLeAspPheLySPhaeAnaBlySThrAlaLysIleThrValAspProLys 540
Db 2253 ATCATTAGAAAGACTTCAAAATTCACACACGCGCATCTCATCTGACTTA 2312
Qy 541 ThrPheGluValPheValAspGlyLySleuCySthSerLySProThrSerGluValPro 560
Db 2313 ACCTTCAGAGCTCTTTAGATGCAAACTCTGACCTTAACCCGCTCTGAAAGTGCT 2372
Qy 561 lEuAlaGlnAArgTyrThrPhePhe 568
Db 2373 CTAGCCAGCGCTACCTTCTTC 2396
RESULT 8
AX356692 2407 bp DNA linear PAT 13-FEB-2002
LOCUS Sequence 10 from Patent EPI176192.
ACCESSION AX356692
VERSION AX356692.1 GI:18674029
KEYWORDS
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Kusters, J.G. and Catcogli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 10 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..2407
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
2..682
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD23261.1"
/db_xref="GI:18674030"
CDS
1..682
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD23261.1"
/db_xref="GI:18674030"
translation="MKLTPEQPKFLYYAGVAAKKRAEGKLNPPEAIATYSAHIM
DFAKRGKTVAMLEBECMHFLKDEVMPEGVNMVDPDVGVEATFPDGTKLVTNWPLEP
DEHFAVGEVFGCDIDIELNAGEVTELENGEPKSLVGHSHFHFPEANLAKLPDRE
KAYGRLDIPSGNTLRIGAGQTRKVLQILPGSKKVIQMGNGLVNNIADBRHKKALEK
AKSHGPIK"
693..2399
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD23262.1"
/db_xref="GI:18674031"
translation="MKMKQEVVNTYPTTGDKVRKLRGDDVLAWEHVDYTYGEELK
GAGKITREGMGQSNPDENTLIDVITNAMIIDYTGKADIGIKNGKIHGIGKAGKD
MODGVPHNVAVGETEALAGEEMIITAGGIDSHTEFLSQEPFLANGVTMFGGT
GVDGDTNATITTPGKWNLRHMLRAAEVSMNVGFLKGNSSKQIVLEIAGALGFK
LIEDGTTSAIDHCLSADEVDOVCITDITVNEAGYDDTLNANGALIAHYIEG
AAGGSPDVTIMAGELNIIIPSSSTPTPIPTINTVAEHLMLMTCHHLDRIRIEDIFS
QSRIRPGSIAADVDHIDIVIAMTSDSQAMGRAGVIRRTQJADKXKKEGKLPED
SADNDNFRIRKIRYISKTTINPALTHGVSIVGEBKADLVVNMVPAFVGKVLK
GAMVNESEMGDSNNAVPTPQVYIREMFGHKKAKPDTSIITVSKVAYVAKGKGL
ERKVLPEVKNCRNIITKKDFKNNKTAHITVDPTFEVFDGKLCSTKPAEVLARVY
PF"

ORIGIN
Alignment Scores:
Pred. No.: 3..31e-187 Length: 2407
Score: 2953.00 Matches: 558

Percent Similarity: 99.12% Conservative: 5
Best Local Similarity: 98.24% Mismatches: 5
Query Match: 98.47% Indels: 0
DB: 6 Gaps: 0
US-09-904-994B-3 (1-568) x AX356692 (1-2407)
Qy 1 MetLysMetLyLeAGlnGluTyrValAsnThrTyrGlyProThrLySLeGlyAspLyVal 20
Db 693 ATGAAAATGAAAAAACAAGAGATGATGTAATACCTTACGACCTACACAGCGCAAAAGTG 752
Qy 21 ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrThyTyrGlyGlu 40
Db 753 CGCTTAGAGAGATACCGATCTTTGGGACAGAAAGTAACATGACTATACCACTTATGGCGAA 812
Qy 41 GluLeuLyPheGlyValAGlyLySerThrIleArgGluGluGlyMetGlyGlnSerSerPro 60
Db 813 GAGCTCAAAATTTGGGCGGGTAAACTATCCGTGAGGCGATGGTCAAGCAATATGTCCA 872
Qy 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
Db 873 GATGAAAACACCCCTTAGATTAGTCATCACACGCGAGATGATTATGACTACACCGGGATT 932
Qy 81 TyrLySaIaAspIleGlyIleLySaengLyLySleHisGlyIleGlyLySaIaGlyAsn 100
Db 933 TACAAGCCGACATTTGGCATTTAAATGCAAAATTCATGCGATTGGCAAGCAGAAAC 992
Qy 101 LySaAspMetGlnAspGlyValSerProHisMetValIleGlyValGlyThrGluAlaLeu 120
Db 993 AAGGACATGACAGAGATGGCGTAAGCCCTCATAGTGTGGGTGGGCAACAAGCATTA 1052
Qy 121 AlaGlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSer 140
Db 1053 GCGGGGAAAGATATGATTTATTCCTGGGGGAGATGATTCACACACCATTCCTCTCT 1112
Qy 141 ProGlnGlnPheProThrAlaLeuAlaAsngLyValThrThrePheGlyGlyThr 160
Db 1113 CCACAACAAATTCCTTACCGCTCTACGCAATGGCGTTTACCAACATGTTGGCGTGACCA 1172
Qy 161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLyThrAsnLeuHisAsg 180
Db 1173 GGTCCGGTAGATGGACGATGCACTACCACTCCGGGGAATGAACTTGACCGCC 1232
Qy 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLySLeGlyAsnSer 200
Db 1233 ATGTTCCGCGACGCTGAAGAGATTTCTATGATGTGGCTTTTGGGCAAGCAATAGC 1292
Qy 201 SerSerLyLySLeGlnLeuValGluGluValGluAlaGlyAlaIleGlyPheLySLeuHis 220
Db 1293 TCCAGTAAAAAACAACCTTGTAACAAATAGAAAGCGGCGCATCGGCTTTTAAATTGCAT 1352
Qy 221 GluAspTrpGlyThrThrProSerAlaIleAspHisGlyLeuSerValAlaAspGluTyr 240
Db 1353 GAAAGCTGGGCGACAACCTCAAGTCAACATCACTCTGTTAGCGTTCAGATGAAATAC 1412
Qy 241 AspValGlnValCysIleHisThrAspThrValaAsngLyValGlyTyrValaAspAspThr 260
Db 1413 GATGCGCAAGTTTGTATTCACACCGATACCGATGAGGCGATTATGTATGATGACACC 1472
Qy 261 lEuAnaAlaMetAsnGlyValArgAlaIleHisAlaTyrHisIleGluGlyValaGlyGly 280
Db 1473 CTGAATGCGATGACGAGGCGCCATCATGCTACCACTGAGGAGCGGCGGAGGAGA 1532
Qy 281 HisSerProAspValIleThrMetAlaGlyGluAsnIleLeuProSerSerThrThr 300
Db 1533 CACTACCTGAGTATATACCAATGACGCGCTCAATATTCACCTCTCCCAACACC 1592
Qy 301 ProThrIleProTyrThrIleAsnThrValaGluHisLeuAspMetLeuMetThrCys 320
Db 1593 CCCACTATCCCTATACCATTTAATACGTTGCGAAGCACTTGAACATGCTCATGACCTGC 1652
Qy 321 HisHisLeuAspLyAsArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340

	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC
A61K39/23,	PC
A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,	PC
C12N1/19,	PC
C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/	PC
53,	PC
G01N33/53, G01N33/566, G01N33/569/(C12N9/80, C12R1:01), (C12Q1/68, PC	
C12R1:01),	PC
C12N15/00, C12N5/00, A61K37/02	CC
Helicobacter felis vaccine	CC
Key Location/Qualifiers	FH
CDS (3)..(683)	FT
(694)..(2181).	FT
Location/Qualifiers	CDS
1..2183	
/organism="Helicobacter felis"	
/mol_type="genomic DNA"	
/db_xref="taxon:214"	
ORIGIN	
Alignment Scores:	
Pred. No.: 5 6e-163	Length: 2183
Score: 2588.00	Matches: 469
Percent Similarity: 99.59%	Conservative: 1
Best Local Similarity: 99.39%	Mismatches: 2
Query Match: 86.30%	Indels: 0
DB: 6	Gaps: 0
US-09-904-994B-3 (1-568) x BD185304 (1-2183)	
OY	1 MetLysMetLysLysGlnGluTyrValAsnThrTYRGlyProThxLysGlyAspLysVal 20
Db	694 ATGAATAATGAAAAAACAAGATGTATTAAACACTCGACCAACCACCAACAGCGCATTAAGTG 753
OY	21 ARGLeUGLYAspTHRAspLeuTPAlaGluVALGIuLHAAPTYrThrThrTYRGlyGU 40
Db	754 CGCTTAGAGAGATACCGCATCTTTGGGCAGAAAGTAAACATGACTATCACCTTAGCGCAA 813
OY	41 GIuLeuLysPhneGlyALagLYLyeThrIleASgGluGlyMetGLyLseAsnSerPro 60
Db	814 GAGCTCAAAATTTGGCGCGGTAAACTATTCCTGAGGGTATGGTCAAGCAATATGCCCA 873
OY	61 ASPGIuAsnThrLeuAspLeuValIleThrAsnAlaMeTlleIlEAsPTYrThrgLyile 80
Db	874 GATGAAAACACCTTAGATTAGTGATCACCAACGGGATATTATGCATCACCGSGGATT 933
OY	81 TYrLysAlaAspLIleGIYLyeLysAnGlyLYSIIehISgLYlEGLYLysAlaGlyAen 100
Db	934 TATTAAGCGCATGTTGATTATAAATGCAAAATCCATGTTATGGCAAGCGGGGAAC 993
OY	101 LysAspMetGlnAspGlyValIsErProHISmetValIGLYValGlyThrgLUalaleu 120
Db	994 AAAAGCAAGCAAGAAGGGGTAAAGCCCTCATATAGTGTGGGTGGGCAAGAAAGCATTA 105
OY	121 ALAglyGluglyMeTlleIleThraLaGlyLYleAspSerTHISthrIspheLusEr 140
Db	1054 GCAGGGGAAGSTATGATTATTTACCGCTGGGGGAGTCGATTGCCAACCACTTCCTCTCT 111
OY	141 ProGInginpheProThraLeuAlaasnGlyValIThrThMeCPheGlyGlyJThr 160
Db	1114 CCCCAACAATTCCTTACCGCTCTACCAATGGTGTATTAACAACATGTTTGAAGGTGCACA 117
OY	161 GLYPROVALAspGlyThraSnAlaThrThrIleThrPRoGlyLYleTPAsnLeuHisAsg 180
Db	1174 GGTCCGGTAGATGGCAACAATGCGACCAACATCACTCCGGGCAATGGAACTTGACCCGC 123
OY	181 MetLeuATGAlaAlaGluGlyTyrSerMetAsnValIGlyPheLeuGlyLYasglyAsnSer 200
Db	1234 ATGTTCGGCGAGCTGAAGAGTATTCTTAGAATGTAGAGGCTTTTGGGCAAAAGGCAATAGT 129
OY	201 SerSerLysLYSgInLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHis 220

|||||
Db TTTAGCAAAAAACAATTGTAGACAGTAGAAGCGGGCGCATTTGGCTTTAAATTGCAT 1353
QY 221 GTUAGPTTGTGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyr 240
Db 1354 GAAGACTGGGGGCAACACCAAGTGGATGCATCTTGAGCGCTGGCAGATGAAATAC 1413
QY 241 AASPValGlnValCysIleHisThrAspThrValAsnGlnAlaGlyTyrValAspAspThr 260
Db 1414 GATGGCAAGTTTGTATCCACCCATACCGGTCAATAGAGCGAGTTATGTGGATGACACC 1473
QY 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValaGlyGly 280
Db 1474 CTAATATGCAATAGACGGGGCGCGCATTCATGCTTACCAATTGAGGAGACGGGGCAGAGA 1533
QY 281 HisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr 300
Db 1534 CACTCACTGATGTTATCCATCGCAGCGGAGCTCAATATTTCTACCCCTCCACACACC 1593
QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
Db 1594 CCCACTATTCCTTATCCATTTATACGGTTGCGAACACTTAGACATGCTCATGACTTGC 1653
QY 321 HisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340
Db 1654 CACCACTGATAGAGCGCATCCGCGAGATTTAACAATTTCTCAAGCCGTATCCGCCCC 1713
QY 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360
Db 1714 GGATCTATTGCGCGCTGAGATGTGCTCCATGATATGGCGTATGCCGATGACTGACTTCC 1773
QY 361 AASPserGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAsp 380
Db 1774 GATTGCGAAGCAATGGGGCGCGCTGGGGAAAGTATCTTAGAACTTGGCAACTGAGAT 1833
QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspArg 400
Db 1834 AAGAATTAATAAAGAAATTTGTAGCTTCTGGAAGATGGTGCAGATACGACAACTTCCGC 1893
QY 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420
Db 1894 ATCAAAACGGTATATCTCCAAATACACCATTAATCCCGCTTTGACCCATGGCGTGAAG 1953
QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhe 440
Db 1954 TATATCGGCTCTGTGGAAGAGGCAAGATCGCGAATTGGTGTGGAACTCTGCTTT 2013
QY 441 PheGlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGly 460
Db 2014 TTGGGCGTAAACCTTAAGATTGATTAAAGGTGGCATGTGGTCTTCTGAAATGGGC 2073
QY 461 AASPserAsnAlaSerValProThrProGlnProValTyrTyrArgLysMetPheGlyHis 480
Db 2074 GATTCTTAAGCGCGCTGCCCGCCAGCTCAGCCGTTTATTAACGCGAAATGTTTGGGAC 2133
QY 481 HisGlyLysAlaLysPheAspThrSerIleThrPhe 492
Db 2134 CACGGCAAGCGCAAAATTTGACACGACATCACTTTT 2169
RESULT 10
AX356689
LOCUS AX356689 2183 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 7 from Patent EP1176192.
ACCESSION AX356689
VERSION AX356689.1 GI:18674026
KEYWORDS
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS Kuster, J.G. and Catolici, G.
TITLE Helicobacter felis vaccine

JOURNAL Patent: EP 1176192-A 7 30-JAN-2002;
Axzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..2183
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
3..5683
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23259.1"
/db_xref="GI:18674027"
/translation="MKLTPKOEKFLYYAGVARKRAEGLKLNQPAIAYISAHIM
DEARGRKTVAEIMEECMPLKDEVMVGVMVDLGVPAFPFGTGLVTVMPLP
DEHFRAGYVKRGCDIDILNCKSTETLELVNNEGRKSLAHVSHFFPFTNKLKPRE
KAYGRKLDIPSGNTLRIGAGGRKQLPLGGSKVKTMGNGLVNNIADBRHKALDK
AKSHGPIK"
694..>2181
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23260.1"
/db_xref="GI:18674028"
/translation="MKMKOEYVNTYPTGDKVRLGDTDMAEVEDHTTYGEBLKF
GAGKTRREGMCGSNSPDENTLDLVTNNMIDYTSIYKADIGIKNGKIHGIGAKND
MODGVSPPMVGVGTREALAGSMITTAGISDHSFPLSPQEPPLANGVTMPFGGT
GPVOSTNATFTYTPGKWNLAHRLRAEEYSNNVGLGKNSKSKQOLVQVEAIGPK
LHEDGTTPSAIDHCLSYADEVDVOCIHHTTVNAGYVDDTLNANGRAIHAHYIBG
AGGGSPPVITMAGELNLPSSSTPTPTVINTVAEHLDMLTCHLDRKREDLQFS
OSRIRPGSIADVEDVLDIGVIAMTSSDSQAGRAEVLPRWTQADKNKREGLPED
GADNDFRIKRYISKYITINDPALVTHSVSYISVEBKIALDVNNMPAFGVGPKIVIK
GMVVFSEMGDSNMSVPTPPQVYVEMGHGKAKFDSIFRVSS"
ORIGIN
Alignment Scores:
Pred. No.: 5,6e-163 Length: 2183
Score: 2588.00 Matches: 489
Percent Similarity: 99.59% Conservative: 1
Beet Local Similarity: 99.39% Mismatches: 2
Query Match: 86.30% Gaps: 0
DB: Gaps: 0
US-09-904-994B-3 (1-568) x AX356689 (1-2183)
QY 1 MetLysMetLysLysGluGlnGlyTyrValAsnThrTyrGlyProThrLysAspLysVal 20
Db 694 ATGAAATGAAATAAACAAGATGATGTAAACACTACCGACCCACACAGCGATTAAGTG 753
QY 21 ArgLeuGlyAspThrAspLeuTyrAlaGluValAlaHisAspTyrThrTyrGlyGlu 40
Db 754 CCTTAGAGAGATACCGATTTTGGGCAGAGTAGAACATGATTAACACTTATGGCGAA 813
QY 41 GlnLeuLysPheGlyValaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro 60
Db 814 GAGCTCAAAATTTGGCGCGGTAAACTATCCCTGAGGATATGGTGCAGAGCAATAGCCCA 873
QY 61 AASPValAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
Db 874 GATGAAACACCTTGAATTAGTATGATCAACCAAGCGAGATATTAATGACATCACCGGAT 933
QY 81 TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100
Db 934 TATTAAGCCGACATTTGTATTAATAATGGCAAAATTCATGATTTGGCAAGCGGGGAAC 993
QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyValaGlyThrGluAlaLeu 120
Db 994 AAAGACATGACAGATGCGGTAGCCCTCATATGTGTGCGGTGGGCGACAGAAACACTTA 1053
QY 121 AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140
Db 1054 GCAAGGGAAGATGATATTAATCCGCTGGGGGAGATCGATTCGACACCACTTCTCTCT 1113

141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr 160
 1114 CCCCAACATTCCTTACCCTCTTACCAATGGTTTACACCAATTTGGAGTGGCA 1173
 161 GlyProValAspGlyThrAsnAlaThrThrLeThrProGlyLysTrpAsnLeuHsArg 180
 1174 GGTCCGGTAGATGGCAAGATGGCAACCATCATCTCCGGGCAATGGAATCTTCACCGC 1233
 181 MetLeuArgAlaAlaGluGlyTrpSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
 1234 ATGTGGGCGCAGCTGAAGATATCTATGATATAGACTTTTGGGCAAGGCAATAGT 1293
 201 SerSerLysLysGlnLeuValGluGlnValGluAlaLeuGlyPheLeuLysHs 220
 1294 TCTACGCAAAAACACTTGTAGAACATGACACCGGCGCGATGGCTTTAAATGGCAT 1353
 221 GluAspTrpGlyThrThrProSerAlaLeuAspHsCysLeuSerValAlaAspGlyLys 240
 1354 GAAGACTGGGGCAACACCAAGTGCATGCATCTGACGCGTGGAGATGAATAC 1413
 241 AspValGlnValCysHsHsThrAspThrValAsnGluAlaGlyTrpValAspAspThr 260
 1414 GATGTGCAAGTTGTATCCACCGATACCGTCAATGACGAGGCTTATGTGATGACACC 1473
 261 LeuAsnAlaMetAsnGlyArgAlaLeuHsAlaLeuHsHsGlyValGlyGly 280
 1474 CTAAATGCAATGAACGGGCGCGCATGCTCAACACATGAGGAGCGGGGAGAGA 1533
 281 HisSerProAspValLeuThrMetAlaGlyLysLeuAsnHsLeuProSerSerThrThr 300
 1534 CACTCAGCTGATGTTATCACCATGGAGGAGAGTCAATATTCACCTCTCTCCACACACC 1593
 301 ProThrLeuProGlyThrThrLeuThrValAlaGluHsHsLeuAspMetLeuMetThrCys 320
 1594 CCCACATTCCTCAATACATTAATACGTTGACAAACATTAATGATGATGACCTGC 1653
 321 HisHsHsLeuAspLysAspArgLeuArgGluAspLeuGlnPheSerGlnSerArgLeuArgPro 340
 1654 CACCACTAGATTAAGCCATCCGCGAGATTTTACAAATTTCTCAAGCCGATATCCGCCCC 1713
 341 GlySerLeuAlaLeuAspValLeuHsAspMetGlyValLeuHsHsThrSerSer 360
 1714 GATCTATTCCTCCGCTGAGATGCTCATGATATGCGCGATGCGGATGACGCTCC 1773
 361 AspSerGlnAlaMetGlyArgAlaGlyGluValLeuProArgThrTrpGlnThrAlaAsp 380
 1774 GATTCGCAACCAATGGGCGCGCTGGGGAAGTGAATTCCTAGAACTTGGCAACTGCAGAT 1833
 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400
 1834 AAGAAATTAATAAAGATTTGGTAACTTCTGAAAGATGTCAGATTAACAACTTCCGC 1893
 401 LeuLysArgTrpLeuSerLysTrpThrLeuAsnProAlaLeuThrHsGlyValSerGlu 420
 1894 ATCAACGCTATATCTCAATACACCATTAATCCGCTTGAACCATGGCGTGAAGAG 1953
 421 TyrLeuGlySerValGluGluGlyLysLeuHsAspLeuValValTrpAsnProAlaPhe 440
 1954 TAAATCGGCTCTCTGGAAAGAGGCGAGATGCCCACTTGGTGGTGAATCTCGCTTT 2013
 441 PheGlyValLysProLysLeuValLeuLysGlyLysMetValValPheSerGluMetGly 460
 2014 TTTGGCGTGAACCTGAATGATGATTAAGTGGCAATGGGCTTCTCTGAAATGGGC 2073
 461 AspSerAsnAlaSerValProThrProGlnProValTyrTrpArgGluMetPheGlyHs 480
 2074 GATTCACGCGCTCGGCGCCAGCGCTCAAGCGGTTTATTAACGCGAAATGTTGGGCAC 2133
 481 HisGlyLysValLysPheAspThrSerThrPhe 492
 2134 CACGGCAAGCGAAATTTGACACGACATCACTTT 2169

RESULT 11

AF330621
 LOCUS 8406 bp DNA linear BCT 13-MAY-2003
 DEFINITION Helicobacter bizozeronii tRNA ribosyltransferase gene, partial cds; GidB, complete cds; urease gene cluster, complete sequence and unknown gene.
 ACCESSION AF330621
 VERSION AF330621
 KEYWORDS AF330621.1 GI:27462193
 ORGANISM Helicobacter bizozeronii
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.
 1 (bases 1 to 8406)
 Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C., McDonough, P., McDonough, S., and Chang, Y.F.
 Cloning and characterization of a Helicobacter bizozeronii urease gene cluster
 DNA Seq. 13 (6), 321-331 (2002)
 22540228
 12652903
 2 (bases 1 to 8406)
 Zhu, J. and Chang, Y.F.
 Direct Submission
 Submitted (21-DEC-2000) Population Medicine and Diagnostic Science, College of Veterinary Medicine, Cornell University, Tower Road, Ithaca, NY 14853, USA
 Location/Qualifiers
 1. 8406
 /organism="Helicobacter bizozeronii"
 /mol_type="genomic DNA"
 /db_xref="taxon:56877"
 <1..929
 /codon_start=3
 /transl_table=11
 /product="tRNA ribosyltransferase"
 /protein_id="AAO15370.1"
 /db_xref="GI:27462194"
 /translation="LMVYERASNKLTHSTFKEIPDPPRHVALVNDTKYIKARLPH TAQRREVEILHRTSPRTCLADMRGKVRGVLGDBGTSCEVLEVDGLGVLSFK HGGELLEMGVLEMLBLGHVPLPYLKRDEMOQDQOVSAKYFGALAAPLASHK FQSADHLLKNFHVETLHVAGCTFLSVOSDIEBHPMHAFCSLPGAGTLDEA PYLCVGTALRATHEHYKRGFGKPELFLHLNPNVXHAQALLTNFHLPESTLIMVA SWVGLDKELXYIATIEHNYRFSYDGMIL"
 926..1465
 /codon_start=1
 /transl_table=11
 /product="GidB"
 /protein_id="AAO15371.1"
 /db_xref="GI:27462195"
 /translation="MRKADLPTHLBMSYVNLGAKKHODIEHNIDSLVDLPI APPKSLDLSGAGPAIPISLACSNAREFLLEPNKRVAFLEHVKALNLNLETOR MRIEHSPOSIVLDLITSRALMAONLIALSAPLEBOGHFLFYKSHLRTETIACDH ECHVYGRVYFYFSQRDPFA"
 1670..2053
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAO15372.1"
 /db_xref="GI:27462196"
 /translation="MLVFGHPOICMPKCVADIOALVPSHTIPFPCHTDAFGL SKHCSDOGKYVAVNQNPLECALMAFKPTIYLAENQLACGSIANDYLDPAIILKYI EHAEOLEBVIKRLDGAIVRSFLML"
 2341..3045
 /codon_start=1
 /transl_table=11
 /product="urease A"
 /protein_id="AAO15373.1"
 /db_xref="GI:27462197"
 /translation="MKLTPELDKMLHAYAGELAKKRKANGKVANTVAVALISAHV BEARPAKGSVADMDGRITLKADVMPGVANHIVEGIEANPDETKVTITPVED GGHKLAGEVILNEDITLNAKQATTLLEVNHGDRPVGVSHFHFEEVVKLEFIRE KAYGRDLIASGTVAVRFEPEKKTVELIQIGGNORLYGFNSLVDRQADTDGKKLAKR AKHNGFGVNCGDKK"

Qy 322 HlaLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
 Db 4027 CACCTGGATTAAGAAACATCAAGAAAGATGTCAGATTGTGATTTAGAGATTGCGCCCA 4086
 Qy 342 SerIleAlaIagLysPheValLeuHlaAspMetGlyValIleAlaMetThrSerSerAsp 361
 Db 4087 ACCATGCGCGCTGAGCAAACTCCAGATATGGGAAATTTCTCATACACAGCTCTGAC 4146
 Qy 362 SerGlnAlaMetGlyValGlnIleValIleProArgThrProGlnThrAlaAspLys 381
 Db 4147 TCCCAACGATGGCGCTGATGAGGATGATCAACCGCATCTTGCAAAACGCGACAAA 4206
 Qy 382 AspLysPheGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAspPheArgIle 401
 Db 4207 AACCAAAAGATTTGGTTCGCTTGAGGAAAAGCGGATTAATGACAACTTCGCGATC 4286
 Qy 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHlaGlyValSerGlyTyr 421
 Db 4267 AAGGCTACATTTCCAAATACACATCAACCCGCTATTGCAACGCGATTTCTGAATAT 4326
 Qy 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhePhe 441
 Db 4327 GTGGGCTCTAGAAAGTGGGAAATGCGCATTTGGGCTTGGAGCTCGCTTCCTT 4386
 Qy 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGlnMetGlyAsp 461
 Db 4387 GGCAATTAAACCAACATGATCATCAAAAGCGGATTCATCGCATTTCTCAATGGCGAT 4446
 Qy 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGlnMetPheGlyHlaIle 481
 Db 4447 GCCATGCTCTATATCCCACTCCCAACCGCTGATTAACCGCAAAATGTTGGCACCAT 4506
 Qy 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501
 Db 4507 GGTAAGCCCAATTTGACCAATATCATCTTTGATCCCAAGTGGCTTATGAAACAGCGC 4566
 Qy 502 ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
 Db 4567 ATTAAAGAGAGTGGCTTGCAAGAGTGTGTTGCAATTAAGTGCAGAACCTGCAGAACCTC 4626
 Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
 Db 4627 ACCAAAAGACCCCAATTCACAGATGTTACCGCACACATCGAAGTCAATCCGAAAC 4686
 Qy 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
 Db 4687 TACAAAGTTAAAGGATGGCAAAAGAGTTACTTCCAAAGCAGCGATTAATACAGCTTA 4746
 Qy 562 AlaGlnArgTyrThrPhePhe 568
 Db 4747 GCACAACTTAACAATTGTC 4767
 RESULT 12
 LOCUS AY295085 1710 bp DNA linear BCT 10-JUN-2003
 DEFINITION Helicobacter pylori urease B (ureb) gene, complete cds.
 ACCESSION AY295085
 VERSION AY295085.1 GI:31580720
 KEYWORDS
 SOURCE Helicobacter pylori
 ORGANISM Helicobacter pylori
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteraceae; Helicobacter.
 1 (bases 1 to 1710)
 The ureb gene sequence of Helicobacter pylori strain isolated from
 central China
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1710)
 AUTHORS Duan, G.C. and Dai, L.P.
 TITLE The ureb gene sequence of Helicobacter pylori strain isolated from
 central China
 JOURNAL Submitted (11-MAY-2003) Epidemiology, College of Public Health,
 Daxue Road 40th, Zhengzhou, Henan 450052, P.R. China

FEATURES
 source
 Location/Qualifiers
 1..1710
 /organism="Helicobacter pylori"
 /mol_type="genomic DNA"
 /strain="MEL-HP27"
 /db_xref="taxon:210"
 /country="China"
 /note="authority: Helicobacter pylori Duan G.C."
 1..1710
 /gene="ureB"
 1..1710
 /gene="ureB"
 /codon_start=1
 /transl_table=1
 /product="urease B"
 /protein_id="AAP51176.1"
 /db_xref="GI:31580721"
 /translation="MKKSRKRVSMVPTTGDKRLGDTDLAEVHDYTYGELK
 FGGKTLRSGMSQNSNPKEDLITNALIVDTYKADIGIKOOKIAGIGGAK
 DMQDGVKNISVGPATLAGEIITYAGIDTHIHIFISPOQIPFASGVTTTGG
 TGPADGVTNATITPPGRNLMRLAEVSMNIGFLAKGNASNDASLADQIEAGIGP
 KIHEDWCTPSALNHLDVADKYDVAIHDTLNEAGVDEDMALTAGMTFTPE
 GAGGAGPDIIKYAGENHILPASTNPTPTVTATEAHMDMLMCHLIDKSIKEDVQF
 ADSRIRPOTIAEDTLHDGIFSTISDSQAMRGVEGVITRTQOTADKNKEFRLKE
 EKDNDNFRIRKYLKTYTNPAIHGISEVSGVGVKADIVLMSAPFGVKNMTI
 KGFRIALSQGDANASIPTPQPYTRMFPHHGAKYDANITTVSKAYIDKIKELG
 LERQVPEVKNCRNITKDMQFNDTNAHIVNPETHYHVFVDSKEVTSKPRKVSIAQLF
 SIF"
 ORIGIN
 Alignment Scores:
 Score: 5.14e-144 Length: 1710
 Percent: 2301.00 Matches: 417
 Percent Similarity: 86.77% Conservative: 75
 Best Local Similarity: 73.54% Mismatch: 75
 Query Match: 76.73% Indels: 0
 DB: 1 Gaps: 0
 US-09-904-994B-3 (1-568) x AY295085 (1-1710)
 Qy 2 LysMetLysAspGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
 Db 7 AAGATTGCAAGAAAGATATGTTCTCATGTATGCGCTTACAGCGCATTAAGTGA 66
 Qy 22 LeuGlyAspThrAspLeuTyrPalaGluValGluHlaAspTyrThrTyrGlyGluGlu 41
 Db 67 TTGGGCGATACAGACTTGATCGCGAAGTGAACATGACATACCATTTATGGCGAAG 126
 Qy 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
 Db 127 CTTAAATTGCGTGGGTAAACCTTGAGAGAAAGGCAATGAGCCATCAACACCTTAC 186
 Qy 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81
 Db 187 AAGGAAGAACTGATTTATATCATCACTAACGCTTTATTCGTGATTCACCGGATTTAT 246
 Qy 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHlaGlyIleGlyLysAlaGlyAsnLys 101
 Db 247 AAAGCGGATTTGGTATTAAGATGGCAAAATCGCTGATGGCAAAAGCGGCAACAA 306
 Qy 102 AspMetGlnAspGlyValSerProHlaMetValValGlyValGlyThrGlnAlaLeuAla 121
 Db 307 GACATGCAAGATGGCTTAAACCAATCTTACGCTGGTCTCTGACTGAAGCCTTAC 366
 Qy 122 GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHlaThrHlaPheLysSerPro 141
 Db 367 GGTGAAGTTTGAATCGTAATGCTGGTGGTATGACACACACATCCACTCATCTCC 426
 Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThrGly 161
 Db 427 CAACCAATCCCTACAGCTTTTGCAAGCGGTGAACAACGATGATGTTGGTGGCGGAAC 486
 Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysThrPheLeuHlaArgMet 181

```

Db      CCTGCGATGGGCACTTAACCAACCACTATCACTCCAGCAGAAAGAAATTTAAATGATG 546
Qy      182 LeuATGAlAAlAGlUGlUTrISerMetAenValGlyPheLeuGlyLySGLyAenSerSer 201
Db      547 CTCAGAGCGCGTGAAGAAATATCTAATGAAATTTAGGTTTCTTAGCTAAAGTAAACGGCTTC 606
Qy      202 SerLySlyGlnLeuValGlnGlnValGlnAlAGlValAlAGlPheLeuSerLeuHISGlu 221
Db      607 AATGATCCGAGCTTACCGCATCAATTAAGCCGGGCGATTTGGCTTTAAATCCATGA 666
Qy      222 AspTrpGlyThrThrProSerAlaIleAspHisCySLeuSerValAlaAspGluTrpAsp 241
Db      667 GACTGGGGAACAACCTCTTCTGCAATCATCATGCGTTAGATGGTCCGACAAATACGAT 726
Qy      242 ValGlnValCysIleHisThrAspThrValAsnGlnAlAGlTyTrValAspAspThrLeu 261
Db      727 GTGCAAGTCGCTATCATCGGACACTTGAATGAAGCCGGTGTGTAGAAAGCACTATG 786
Qy      262 AsnAlaMetAsnGlyAArgAlaIleHisAlaTyHisIleGlnGlyValAGlGlyGlyHis 281
Db      787 GGAGCCATTCGCCGACGACCTATGACACTTTCACACTGAAGCGCTGCGGAC 846
Qy      282 SerProAspValIleThrMetAlAGlGlyGlnLeuAsnIleLeuProSerSerThrTrpPro 301
Db      847 GCTCCGATATCATTTAAAGTAAAGTCGCGCGGACCAACATTCGCGCTTCCACTAACCC 906
Qy      302 ThrIleProTyThrIleAsnThrValAlAGlHisLeuAspMetLeuMetCysHis 321
Db      907 ACTATCCCTTCTACTGTGAATACAGAAACACATGACACTGTCTTAGTGTCAC 966
Qy      322 HisLeuAspLyAArgIleArgGlnAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db      967 CACTTGATTAAGAAGCTTTAAAGAAATGTCAGTTCGCTGATTCAGATCCGCCCTCAA 1026
Qy      342 SerIleAlaAlAGlUAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
Db      1027 ACCATTCGCGGTGAAGACACTTTCATGACATGACGCGGATTTTCTCATCACTAGTTCGAC 1086
Qy      362 SerGlnAlaMetGlyAArgAlAGlGlyValIleProArgThrTrpGlnThrAlaAspLyS 381
Db      1087 TCTCAAGCTATAGGTCGTGTGGTGAAGTATACACAGAACTTGGCAACAGCTGCAAA 1146
Qy      382 AsnLySlyGlnPheGlyLySLeuProGluAspGlyLyAspAsnAspAsnPheArgIle 401
Db      1147 AACCAAAAAAGATTTGGCGCTTGAAGAAGAAAGGCGATTAACGACAACTTCAGAAATC 1206
Qy      402 LysArgTyrlSerLySlyThrIleAsnProAlaLeuThrHisGlyValSerGluTy 421
Db      1207 AAACGCTACTGTGCTAAATACACCATTAACCCAGCATCGCTCATGAGATTAGCGAGTAT 1266
Qy      422 IleGlySerValGlnGlnGlyLySleAlaAspLeuValValTrpAsnProAlaPhePhe 441
Db      1267 GTAGGTTCTGTAGAAAGTGGCAAAAGTGGTGAAGTGTGTATGTGGAGTCCAGATTCCTT 1326
Qy      442 GlyValLyProLySleValIleLySGLyGlyMetValValPheSerGlnMetGlyAsp 461
Db      1327 GCGCGTAAACCCAACTGATCATCAAGGTGGTTATTCATGATGATCAATGGCCGAT 1386
Qy      462 SerAsnAlaSerValProThrProGlnProValTyTrArgGlnMetPheGlyHisHis 481
Db      1387 GCGAAGCTTATCTACCTACCCCAACAGAGTTTATTAAGAAATGTTCGCTCATCAT 1446
Qy      482 GlyValAlaLysPheAspThrSerIleThrPheValSerLySlyValAlaTyTrGluAsnGly 501
Db      1447 GGTAAGCCCAAAATACATCAACATCACTTTTGTCTTAAAGCGGCTTATGCAAAAGGC 1506
Qy      502 ValLySGLyLysLeuGlyLeuGlnValArgGlnValIleProValLyAsnCysArgAsnIle 521
Db      1507 ATTAAGAAAGATTTAGGCTTGAAGACAAAGTTCCTCCGTTAAAAATTGCGAAGAACATC 1566
Qy      522 ThrLySlyAspPheLysPheAsnAspLySThrAlaLySleThrValAspProLySThr 541

```

```

Db      1567 ACTAAAAAGACATCAATTCAACGACACTACCGCTCACTTGAAGTCAATCTGAAACT 1626
Qy      542 PheGlnValPheValAspGlyLySLeuCySThrSerLySProThrSerGlnValProLeu 561
Db      1627 TACCAATGTTCTGTAGTGAAGGCAAGAAAGTAACTTCTAAACGACCACTAAAGTGA 1686
Qy      562 AlaGlnArgTyThrPhePhe 568
Db      1687 GCGCAACTCTTATGACATTTTC 1707

RESULT 13
HECUREASE      2664 bp      DNA      linear      BCT 03-MAY-1995
LOCUS          Helicobacter heilmannii urease, complete cds'.s.
DEFINITION     L25079
ACCESSION      L25079
VERSION        L25079.1 GI:793908
KEYWORDS       urease.
SOURCE         Helicobacter heilmannii
ORGANISM       Helicobacter heilmannii
                Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
                Helicobacteraceae; Helicobacter.
REFERENCE      1 (bases 1 to 2664)
                Solnick,J.V., O'Rourke,J., Lee,A. and Tompkins,L.S.
                Molecular analysis of urease genes from a newly identified
                uncultured species of Helicobacter
                Infect. Immun. 62 (5), 1631-1638 (1994)
JOURNAL        9422523
MEDLINE        8168924
PUBMED         On May 3, 1995 this sequence version replaced gi:529422.
                Original source text: Helicobacter heilmannii (individual isolate
                2) DNA.

FEATURES
    source
    1..2664
        /organism="Helicobacter heilmannii"
        /mol_type="genomic DNA"
        /isolate="2"
        /db_xref="taxon:35817"
        /lab_host="BALB/c"
        200..204
        211..915
        /codon_start=1
        /transl_table=11
        /product="urease"
        /protein_id="AA065722.1"
        /db_xref="GI:529423"
        /translation="MKLRPEKLDKMLHYAGELAKORAKGKIKLYTEAVALLISAWM
        EBARKGKSVADLNOBGRITLKADVDMEGVAMHIEVGI EAGFPDGTGLVTIHTPEVA
        GSDKLAPSEVILKNEIDITLNGKAVOLKVKKKGRPVQVCSHFPEVNVKLLDPDR
        KAYGRLLDASGTAVRFPGRBKTVELIDIGNKRIVGFNLVDRQADHDGKQALKR
        AKERKFTINGCDNRK"
        919..923
        930..2636
        /codon_start=1
        /transl_table=11
        /product="urease"
        /protein_id="AA065723.1"
        /db_xref="GI:529424"
        /translation="MKKTSRKEYVSMGPPTGDKVRLGDDTLILEVHDCITTYGEIK
        FEGGKTRIDMGQINSBSHELDVITNALVDVTGYTKADITGKNGKIRIGKRGKNGK
        DLQDVCNRLCVGPATELAAEGLVITAGIDITHLHFSPOQIPAFASGLITTMIGG
        TLPADGTAAATITPGRNMLKEMLRASEBEYANMLGLGKGVNSFEPALIDOLEAGAGF
        KIHEDMGSTPEALINALNIADKYDVQVALIHDTLNAECVDELTLEALIDQRTIHTPTE
        GAGGHAARDVIMAGEFNILPASTNPTI PFTKNTAEHMDMLMCHLNLKIKEDVEFA
        DSRIRPQTIADKDKHMDGISITSSDQAMGRVGEVITRTMOTADKKKKEGRLPEE
        KQDNDFRIKIKYIKTITNPALITGISTVSGSEVGYKADLVLMSPAFGKIKPMNLIK
        GGFALISMGKANASI PTPQPVYKREMGHNGKAFDNIITFVSQVAENGIKHEILG
        QRVVLPRVKNCRNITKDKLFNDVTAHIEVNPETYKVKVDGNEVTSAAADKLSLAOLYN
        LP"

ORIGIN
Alignment Scores:
Pred. No.:      1..296-143      Length:      2664
Score:          2298.50      Matches:      418

```

Percent Similarity: 86.77% Conservative: 74
 Best Local Similarity: 73.72% Mismatches: 74
 Query Match: 76.64% Indels: 1
 DB: 1 Gaps: 1

US-09-904-994b-3 (1-568) x HECUREASE (1-2664)

```

QY      2  LyeMetLysylserGlnGluYrValAsnThrYrGlyProThrLysGlyAspLysValArg 21
DB      936  AAAAATTTCTGAAAAAGAAATGTTTCTATGATGACCCACTACGGCGGATAAAGTCCGC 995
QY      22  LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspYrThrThrYrGlyGlu 41
DB      996  TTGGGCGACACTGACCTGATCCTTAAAGTGCAGAACATGACACACTTATGGCGAAAGA 1055
QY      42  LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
DB      1056  ATCAAAATTCGGGGCGGAAAAACCATCCGCGATGGGACAACCAACGACGCCAGC 1115
QY      62  GluAsnThrLeuSerLysValIleThrAsnAlaMetIleLeuAspYrThrGlyIleTyr 81
DB      1116  AGCCATAGCTCGATCTTGTCATCACCAACGCGCTGATGTGATTCACCGGCAATTAC 1175
QY      82  LysAlaAspIleGlyIleLysAsnGlyLysIleGlyIleGlyLysAlaGlyAsnLys 101
DB      1176  AAAGCCGACATTGGCATTTAAAGCGCAAAATCCACGGGATGGCGAAAGCGCGACAA 1235
QY      102  AspMetGlnAspGlyValSerProHisMetValGlyValGlyIleThrGluAlaLeuAla 121
DB      1236  GACCTACAAATGGCGTTTGCACAGGCTCTGTAGACCTGTAGACCTGTAGACGCTGCT 1295
QY      122  GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
DB      1296  GCTGAAGCTTGATGTGTACCGCGGTGATGCACACCAACATCCATTCATTTCTCCT 1355
QY      142  GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyIleThrGly 161
DB      1356  CAACAATCCCTTACTGCTTTGCTAGCGGATCCACCATGATCGGGGCGGACATCGA 1415
QY      162  ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysLysPheLeuHisAspMet 181
DB      1416  CTTGCAGATGGACCAACGCGCACCATCATCTCCGCGCTGGAATTTAAAGAAATG 1475
QY      182  LeuAlaGlyAlaGluGlyYrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB      1476  CTCGGGCTTCTGAAGATAGCGCATGAACCTTGCTTACCTTGTTAAGGGGAATGTCT 1535
QY      202  SerLysLysGlnLeuValGlnGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGly 221
DB      1536  TTTCGAACCTGCTCTCATTTGACAGCTGAAAGCCGCGCGATGGCTTTAAATCCAGAA 1595
QY      222  AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTrpAsp 241
DB      1596  GACTGGGGGACACACCTCAGCATCAACACGCTCTTAAACCTGCTGACAAATACGAT 1655
QY      242  ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyYrValAspAspThrLeu 261
DB      1656  GTGCAATGGCGGACGACGACGACCTTGAATTAACCGCGCTGTGGGAAGAACCTTA 1715
QY      262  AsnAlaMetAsnGlyAlaGlyAlaIleHisAlaTyrHisIleGluGlyAlaGlyHis 281
DB      1716  GAAGCCATCGCTGACGACGACGACCTTCCACACCGAAGTGTGGGCGGAGACAC 1775
QY      282  SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThrPro 301
DB      1776  GCTCTGACCTGATCAAAATAGCGGCGGAATTTAACTCTTCTGCTTACCAACCCC 1835
QY      302  ThrIleProYrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetCysHis 321
DB      1836  ACCATCCCTTTACCAAAACACGAGACCGGAACACATGACAGCATGATTATG--TGCAC 1892
QY      322  HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB      322  HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341

```

```

DB      1893  CACTTGGATATAAATCATCAAAAGATGTGAAATTTGCCGACTCAGCTACGCCGCCAA 1952
QY      342  SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB      1953  ACCATTGCGCGGAAAGATTAATCTCCAGCATGGGAGTTTCTTCATCATCAGTACCTCCGAC 2012
QY      362  SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
DB      2013  TCTCAAGCCATGGGTCCGCGGGCGAAGTATACCCGCACTTGGCAACACGCGGACAA 2072
QY      382  AsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspArgIle 401
DB      2073  AACAAAAGAAATTTGGCCGCTTGCCTGAGAAAGGCGACACACGCAACTTCCGATC 2132
QY      402  LysArgTyrIleSerLysYrThrIleAsnProAlaLeuThrHisGlyValSerGlyTyr 421
DB      2133  AAACGCTATTTTCCAAATACACATCAACCTCCCATCACACACGCGCATTTCTGAATAT 2192
QY      422  IleGlySerValGluGluGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhePhe 441
DB      2193  GTGGCTCTGAGAGGTGGGTAAATACGTGACTGTGCTTGGAGCCCTGGGCTCTT 2252
QY      442  GlyValLysProLysIleValIleLysGlyGlyMetValAlaPheSerGluMetGlyAsp 461
DB      2253  GGCAATTAAACCAACATGATCATCAAGCGGTTTCACTTGGCGTTCTCAATGGCGCAT 2312
QY      462  SerAsnAlaSerValProThrProGlnProValIleTyrArgGluMetPheGlyHisHis 481
DB      2313  GCCAACCTTCTATCCCACTCCCAACCCGTTACTACCGCGCAAAATGTTCGCGCACAC 2372
QY      482  GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValIleTyrGluAsnGly 501
DB      2373  GGCAGAGCCAAATTTGACCAACATCATCTTGTGTCTGAAGGGCTTATGAAAGGCG 2432
QY      502  ValLysGluLysLeuGlyLeuGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
DB      2433  ATTAAACAGAGTTGGCTTGCAAGAGTGTGTGCTGCTGTGAAGAACTCCGCAATC 2492
QY      522  ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
DB      2493  ACCAAAAGAACTTAAGTTCAACGATGTCCACCGCACATCGAAGTCACACCTGAAACC 2552
QY      542  PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
DB      2553  TACAAAGTGAAGGATGGACCAAGATTACCTCCATCGGCTGACAAATGAGCTTA 2612
QY      562  AlaGlnArgYrThrPhePhe 568
DB      2613  GCACAACTCTATAACCTATTC 2633

```

RESULT 14
 A07398 LOCUS 1710 bp DNA linear PAT 20-JUL-1993
 DEFINITION C.jejuni gene for 61 kDa protein.
 ACCESSION A07398
 VERSION A07398.1 GI:412236
 KEYWORDS
 SOURCE Campylobacter jejuni
 ORGANISM Campylobacter jejuni
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 Campylobacteriaceae; Campylobacter.
 1 (bases 1 to 1710)
 Labigne A.
 Nucleotide sequence codings for a protein with urea reactivity
 Patent: EP 0367644-A 8 09-MAY-1990;
 PASTEUR; INSERM
 INSTITUT location/Qualifiers
 1. 1710
 /organism="Campylobacter jejuni"
 /mol_type="unassigned DNA"
 /db_xref="taxon:197"
 1. 1710
 /codon_start=1

CDS

```

/transl table=11
/product="61 kDa protein"
/protein_id="CAA00663.1"
/db_xref="GI:41237"
/db_xref="GOA:P14917"
/db_xref="UniProt/Swiss-Prot:P14917"
/translation="MKKSRKEYVSMYGGTGDKRYLAGEPTDLIAVEHDYTIYGEELK
FGGGTIRREGMSQNNPSKEELDIITNALIVDTYGIYKADIGIDGKTAGIAGKGNK
DMQDVKNNLSVGPRTAELALAGSLVTVAGITDTHIHFIISPOQIPPAFASGVTMIGCG
TTPADGTNATTTTPGRRLKMLRAEYSISNMGFLAKGNASNDLADQIEAGALGF
KIHEDWGTTPSAIINHALDVADKYDQVALIHDTLNEAGCVEDTMAIARITMTPTHE
GAGGGHAPDIIRVAGEHNILPASTNPTIPFTVTELEHMDMLVCHILDKSJKEDVOF
ASRIRPOTIAEDTLHDWGIFFSTSSDQAMGRVEVETRTQTADKDKKFGKGLKE
EKGDNDNFRILKRYLSKTYTNPAIHGISISYGVSEVGYKADVLVMSPAFGKPMII
KCGFIASOMGDANASITPTPOVYVREMHAKGAKYADNATIPVSGAADKIKREIG
LERQVLPVKNCRNITKDMQFNDTTHAIEVNPBTYHVPDGEVTSKPAKVSIAOLF
SIF"

```

ORIGIN

Alignment Scores:

Pred. No.:	9_48e-144	Length:	1710
Score:	2297.00	Matches:	416
Percent Similarity:	86.77%	Conservative:	76
Best Local Similarity:	73.37%	Mismatches:	75
Query Match:	76.59%	Indels:	0
DB:	6	Gaps:	0

US-09-904-994B-3 (1-568) x A07398 (1-1710)

```

QY 2 LysMetLysLysGlnGlnLysValAsnThrTYRGLYProThrLysGlyAspLysValArg 21
DB 7 AAGATTAGCAAAAGAAATATGTTTCTATGATGGTCCTACTACAGCGCAATAAGTGAGA 66
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTYRThrTYRGLYGlnGlu 41
DB 67 TTGGCGGATACAGACTGTATCGCTGAAGTAGAACATGACTACCCATTATGCGCGAAGAG 126
QY 42 LeuLysPheGlyAlaGlyLysThrIleATRGluGlyMetGlyGlnSerAsnSerProArg 61
DB 127 CTTAAATTTGGTGCGGCTTAAACCTTAAGAGAGGAGATGAGCCAACTTAAACACCTTAG 186
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTYRThrGlyIleTYR 81
DB 187 AAAGAAAGAGTGGATTATTAATCACTTAACGCTTAATCGTGATTACACCGGTATTAT 246
QY 82 LysAlaAspTleGlyLysAsnGlyLysIleHisGlyTleGlyLysAlaGlyAsnLys 101
DB 247 AAAGCGGATATGTGTATTAAGATGCAAAATCGCTGCAATTGTGTAAAGCGGTAAACAA 306
QY 102 AspMetGluAspGlyValSerProHisMetValIleGlyValGlyThrGlnAlaLeuAla 121
DB 307 GACATCCAAAGATGGCTTAAAAACAATCTTAGGTAGGTCCTGCTACTGAAGCCTTAGCC 366
QY 122 GlyGlnGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
DB 367 GGGAAAGCTTTGATCGTAACGGCTGGTGATTAACAACAATCCACTTCATTCACTCACCC 426
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrThrMetPheGlyGlyGlyThrGly 161
DB 427 CAACAAATCCCTTACACCTTTTGCAGACGGGTGTAACAACATGATTTGGTGGTGAACCGGT 486
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
DB 487 CTTGCTGATGGCACTTAATCGACTACTTCACTCCAGGCAAGAAATTTTAAATGGAATG 546
QY 182 LeuArgAlaAlaGlyLysLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB 547 CTCAGAGCGCGCTGAAGAATATTTCTATGAATTAAGTTTCTTGGCTAAAGTAAACGTTCT 606
QY 202 SerLysLysGlnLeuValGlnGlnValAlaGlyAlaIleGlyPheLysLeuHisGly 221
DB 607 AAGATGCGAGCTTACCGATCAAAATTGAACCGGCGGTGCAATTGGCTTTAAATTCACGAA 666

```

```

QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyLysArg 241
DB 667 GACTGGGGCACCACTCTTCTTGCATCAATCATGAGATGTTGGCGGACAAATACGAT 726
QY 242 ValGlnValCysIleHisThrAspThrValaGlnGluAlaGlyTYRValaAspAspThrLeu 261
DB 727 GTGCAAGTCGCTATCACACAGACACTTGAATGAAGCCGGTGTGTGAACAACACTAAG 786
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTYRHisIleGlnGlyValGlyGlyHis 281
DB 787 GCTGCTATTTCTGAGAGCACTAAGCACTTCCACACTGAAGGCGCTGGCGGAGCAC 846
QY 282 SerProAspValIleThrMetAlaGlyGlnLeuAsnIleLeuProSerSerThrThrPro 301
DB 847 GCTCTGATATTTATTAAGATAGCCGGTGAACAACAATCTTCCGGCTTCCACTAACCC 906
QY 302 ThrIleProTYRThrIleAsnThrValAlaGluHisLeuAspMetLeuMetCysHis 321
DB 907 ACCATCCCTTTCACCGTGAATACAGAACACAGACACATGACATGCTTATGCTGTGCCAC 966
QY 322 HisLeuAspLysArgIleArgGlnAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 967 CACTTGATTAAGAAAGCAATTAAGAAAGATGTCAGTTGCTGATTCAAGATCCGCCCTCAA 1026
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 1027 ACCATTGGCGGTGAAGAGACACTTGTGATGACATGGGGATTTTCTCATACCAAGTTCTGAC 1086
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
DB 1087 TCTCAAGCGATGGCGCGTGGGTGAAGATTATCACTAGAACTGGCAACACCTGACAAA 1146
QY 382 AsnLysLysGlnPheGlyLysLeuProGlnLysAspAsnAspAsnPheArgIle 401
DB 1147 AACCAAGAAAGATTGGCGCGCTTGAAGAGAAAAGCGGATTAAGCAACTTCACGATC 1206
QY 402 LysArgTYRLeSerLysTYRThrIleAsnProAlaLeuThrHisGlyValaSerGlyTYR 421
DB 1207 AAACCTACTTCTTAATATACCATTAACCCAGGAGATCCGCTCAAGGATTAAGCGATAT 1266
QY 422 IleGlySerValGlnGlnGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhePhe 441
DB 1267 GTAGGTCAAGTAGAAGTGGGCAAGGTGCTGACTTGTGATGTGAGAGTCCACACTTCTT 1326
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValIlePheSerGlyMetGlyAsp 461
DB 1327 GCGTGAACCCCAACATGATCATCAAGCGGATTCATTCGCTTAAGCCAAATGGCGAT 1386
QY 462 SerAsnAlaSerValProThrProGlnProValTYRTYRArgGluMetPheGlyHisHis 481
DB 1387 GCGAAGCTTCTATCCCTACCCCAACAACCGGTTTATTAAGAAATGTGTGCTCATCAT 1446
QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTYRGlusnGly 501
DB 1447 GGTAAAGCTAAATAGAGTGAACAACTCACTTGTGTCTCAAGCGGCTTATACAAAGGC 1506
QY 502 ValLysGlnLysLeuGlyLysGlnLysArgGlnValaLeuProValLysAsnCysArgAsnIle 521
DB 1507 ATTAAGAAAGATTAGACTTGAAGACAAGTGTGGGTAATAAATTCGAAATATATC 1566
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValaAspProLysThr 541
DB 1567 ACTAATAAAGAACATGACATTCACACACTACTGCTCATATTGAAGTCAATCTGAACCT 1626
QY 542 PheGlnValaPheValaAspGlyLysLysLeuCYThrSerLysProThrSerGlnValaProLeu 561
DB 1627 TACCATGTGTTGTGATGCGCAAGAAAGTAACTTAAACCGACCAATAAAGTGAAGCTTG 1686
QY 562 AlaGlnArgTYRThrPhePhe 568
DB 1687 GCGCAACTCTTAAATTTTC 1707

```

RESULT 15

A08818
LOCUS A08818 1710 bp DNA linear PAT 09-AUG-1993
DEFINITION C.jejuni DNA for 61 kDa protein.
ACCESSION A08818
VERSION A08818.1 GI:412247
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Campylobacteraceae; Campylobacter.
REFERENCE
AUTHORS
TITLE
SEQUENCES OF NUCLEOTIDES CODING FOR A PROTEIN HAVING AN UREASIC
ACTIVITY
JOURNAL
FEATURES
source
PATENT: WO 9004030-A 8 19-APR-1990;
Location/Qualifiers
1. 1710
/organism="Campylobacter jejuni"
/mol_type="unassigned DNA"
/db_xref="taxon:197"
1. 1710
/codon_start=1
/transl_table=1
/product="61 kDa protein"
/protein_id="CA00811.1"
/db_xref="GI:412248"
/db_xref="GOA:P14917"
/db_xref="UniProt/Swiss-Prot:P14917"
/translation="MKKSRKEVSMYPTGDCVRLGDDLLAENEHDYTIYGEELK
FGGKTLREGSGSNPSKEELDLITNALIVDTYGYKDIGKDKKIGIGGAK
DMQGVKNNTSVGPATLAGEGLIYTAGIDPHIHFIHQPIPTFASGVTTMIGG
TGADGTNATTTTPGRNLKMLRAEYSNLCFLKAGNSDASLDQIEAGATP
KIHEDWCTTSALNHADVADKYDQVAIHDTLNEAGCVEDMVAALAGRTMTEFE
GAGGAPDIIKYAGEHNIIPASTNPITPTVTEAEHMDLWCHLIDKSIKEDVF
ADGRIPQTLAEVTLDMGIFSTSSDSQAMGVGIVRTWQADKNKEFGELKE
EKGDNDNFRIKRYLKYTINPAIHAGISEVGVKADLVIMSPAFKPKVNMII
KGGFIALSOMGDANASTPTPOPVYRPMFPAHAKAKDANITTPSQAAVDFKIGKEIG
LERQVLPVKRCRNITKQWQFNDDTAHIEVNPETIYHVFVDEKEVTSRKANKVSLAQLF
SIP"

ORIGIN
Alignment Scores:
Pred. No.: 9.48e-144 Length: 1710
Score: 2297.00 Matches: 416
Percent Similarity: 86.77% Conservative: 76
Best Local Similarity: 73.37% Mismatches: 75
Query Match: 76.59% Indels: 0
Gaps: 0
DB: 6
US-09-904-994B-3 (1-568) x A08818 (1-1710)

QY 2 LysMetLysLeuSgInGluTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
Db 7 AAGATTGACGAAAGAAATATGTTCTATGATGTCTCTACAGCGCAATAAGGTAGA 66
QY 22 LeuGlyAspThrAspLeuTTPAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41
Db 67 TTGGGCGATACAGACTTGATCGCTGGAAGTAGACATACACCATTTATGGCGAAGAG 126
QY 42 LeuLysPheGlyAlaGlyLysThrTleArgGluGlyMetGlyGlnSerAspSerProAsp 61
Db 127 CTTAAATTCGGTGGCGGTAAACCCCTAAGAGAGAGCATGAGCCATCTTAACAAACCCCTAGC 186
QY 62 GluAsnThrLeuAspLeuValIleThrAspAlaMetCileIleAspTyrThrGlyIleTyr 81
Db 187 AAGAAGAGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 246
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
Db 247 AAGCGGATTTGGTATTAAGATGCGCAAAATCCGTCGCAATGGTAAAGCGGTAAACAA 306
QY 102 AspMetGlnAspGlyValIleSerProHisMetValGlyValGlyTyrThrGluAlaLeuAla 121

Db 307 GACATGCAAGATGCGCTTAAACAACTTAAGCTAGTCTGCTACTAGCAAGCTTAGCC 366
QY 122 GlyGluGlyMetCileIleThrAlaGlyValIleAspSerHisThrHisPheLeuSerPro 141
Db 367 GGTGAAGTTTGATTCGTAAAGCGCGTGTGTATGACACACACATCCCTCATATTGACCC 426
QY 142 GlnIlePheProThrAlaLeuAlaAsnGlyValThrHisMetPheGlyGlyValThrGly 161
Db 427 CAACAAATCCCTTACAGCTTTTGGACAGGTGAACAAACATGATGTTGGTGGTGAACCGGT 486
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTyrAsnLeuHisArgMet 181
Db 487 CTGTCGATGCGCACTAATGCGATGACATGATCATGTCGCGAGAGAAATTTAAATGATGAT 546
QY 182 LeuArgAlaAlaGluGlyTyrSerMetAspValGlyPheLeuGlyLysGlyAsnSerSer 201
Db 547 CTCAGACGGCTGAAAGAAATTCATGAAATTGATGTTCTTGGCGTAAAGCGTTCTT 606
QY 202 SerTyrLysGlnLeuValGluGlnValGluAlaGlyValAlaIleGlyPheLysLeuHisGlu 221
Db 607 AACGATGCGAGCTTAGCGCATCAAAATGAAAGCGGTGCGATGCGCTTAAATTCACGAA 666
QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241
Db 667 GACTGGGGCACCACTCTTCTGCAATCAATCAATGATGTTGGCGCAAAATACGAT 726
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261
Db 727 GTGCAATCCGCTATCCACACAGACACTTGAATGAAGCCGTTGTGTAAGAACACATATG 786
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisGlyIleGlyValGlyGlyHis 281
Db 787 GCTGCTATTCCTGACCGCACTATGACACTTTCCACACTAAGCGGTGCGGCGACAC 846
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301
Db 847 GCTCTGATATTATTAAGTAGAGCCGGGAAACAAACATTTCTCCGCTTCCACTAAGCCC 906
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
Db 907 ACATCCCTTCCACCGGTAATACAGAAAGACAGACACATGACATGCTTAGTGTGTCAC 966
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 967 CACTGTGATTAAGCATTAAGAAAGATGTCAGTTCCCTATTCAAGATCCCGCTCAA 1026
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
Db 1027 ACCATTGCGGCTGAAGACACTTTCATGATGATGGGATTTTTCATACACCAATTCAGAC 1086
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381
Db 1087 TTCACAGCATGAGCGCGGTGGGGAAGTATACATCACTGAACCTTGGCAACAGCTGACAA 1146
QY 382 AsnLysLeuGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
Db 1147 AACCAAGAAATTTGGCCGCTTGAAGAAAGGCGCATGACGAACTTCAGAGATC 1206
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGluTyr 421
Db 1207 AAAGCTACTTGTCTTAATACACCATTAACCCGACGATCCCTAGATGAGATTAAGCAATAT 1266
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValTyrAsnProAlaPhePhe 441
Db 1267 GTAGGTTCACTAGAGAGGCGCAAGCGCGTGAATGATGATGATGATGATGATGATGAT 1326
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
Db 1327 GGGGTGAACCAACATGATCAACAAAGCGGATTCATTCGTTAAGCCAAATGGGGAT 1386
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHisHis 481
Db 1387 GCGAAGCTTCTATCCCTACCCCAACCGGTTTATTAAGAAATGTTCCGCTCATCAT 1446

Qy 482 GlyValAlaIysPheAspThrSerIleThrPheValSerIysValAlaTyrGluAsnGly 501
Db 1447 GGTAAAGCTAAATACATGCACAAACATCCTTTGTGTCTCAAGCGCTTATGCAAAAGC 1506
Qy 502 ValLysGluLysLeuGlyLeuGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
Db 1507 ATTAAGAAGAAATTAGGACTTGAAAGACAAGTGTTCGCGTAAATAATTGCAGAAATATC 1566
Qy 522 ThrLysLysAspPheLysPheAsnAspIleThrAlaLysIleThrValAspProLysThr 541
Db 1567 ACTAAAAAAGACATGCATTCAACGACACTACTGCTCACATTGAAGTCAATCCTGAAACT 1626
Qy 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
Db 1627 TACCATGTGTGTGATGCGCAAGAGTAACCTCTAAACCAGCCAATAAAGTGAAGCTTG 1686
Qy 562 AlaGlnArgTyrThrPhePhe 568
Db 1687 GCGCAACTCTTTAGCATTTTC 1707

Search completed: February 15, 2005, 23:08:46
Job time : 5478.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_p1us_p2n model

Run on: February 15, 2005, 20:44:36 ; Search time 220.332 Seconds
(without alignments)
4218.194 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKKQEVYVNTGPTKGDV.....KLCTSKPTSGVPLAQRVTF 568

Scoring table:
BL0SUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 81813359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09904994/runat_11022005_145648_25560/app_query.fasta_1.1102
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdd
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p1us -NORM=ext HEADSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USRR=US09904994.@CGN_1.1.77.@runat_11022005_145648_25560 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/6C.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297	76.6	2735	2	US-08-920-095-1
2	2297	76.6	2735	5	PCT-US96-05800-1
3	2281	76.1	4824	4	US-09-431-705-1
4	2281	76.1	4824	4	US-09-431-705-1
5	2244	74.8	2619	2	US-08-467-822-19
6	2244	74.8	2619	3	US-08-432-697-19
7	2244	74.8	2619	3	US-08-466-248-19
8	2028.5	67.6	6131	1	US-07-732-242C-8
9	1828	61.0	1701	4	US-09-252-991A-10316
10	1826	60.9	1716	5	US-08-487-429A-3
11	1826	60.9	1716	5	PCT-US96-05320A-541
12	1826	60.9	1830121	4	US-09-557-884-1

13	1826	60.9	1830121	4	US-09-643-990A-1	Sequence 1, Appl
14	1821	60.7	1878	4	US-09-489-039A-2045	Sequence 2045, Ap
15	1818.5	60.6	1710	4	US-09-543-661A-1857	Sequence 1857, Ap
16	1818	60.6	1704	4	US-09-328-352-1786	Sequence 1786, Ap
17	1817	60.6	2400	1	US-08-967-513-1	Sequence 1, Appl
18	1817	60.6	2400	2	US-08-687-645B-1	Sequence 57, Appl
19	1809	60.3	87563	3	US-09-453-702B-57	Sequence 256, App
20	1808	60.3	8729	3	US-09-453-702B-258	Sequence 117, App
21	1789.5	59.7	1716	4	US-09-710-279-3828	Sequence 3828, Ap
22	1789.5	59.7	3164	4	US-09-710-279-3828	Sequence 3853, Ap
23	1789.5	59.7	3234	4	US-09-710-279-3853	Sequence 2189, Ap
24	1788.5	59.6	1722	3	US-09-134-001C-2189	Sequence 22, Appl
25	1784.5	59.5	5966	4	US-08-956-111E-22	Sequence 10546, A
26	1784.5	59.5	5966	4	US-08-781-986A-22	Sequence 13, Appl
27	1761	58.7	1653	4	US-09-252-991A-10546	Sequence 89, Appl
28	1708	57.0	1833	4	US-09-602-777A-13	Sequence 15, Appl
29	1626.5	54.2	1797	4	US-09-601-198-89	Sequence 2, Appl
30	1619	54.0	1625	4	US-09-602-777A-15	Sequence 10513, A
31	1610	53.7	4403765	3	US-09-103-840A-2	Sequence 4032, Ap
32	1610	53.7	4411529	3	US-09-103-840A-1	Sequence 1053, A
33	1115	37.2	951	4	US-09-252-991A-10513	Sequence 121, App
34	945.5	31.5	3475	4	US-09-710-279-4032	Sequence 7157, Ap
35	701	23.4	636	4	US-09-252-991A-10653	Sequence 4755, Ap
36	317	10.6	309	4	US-09-710-279-121	Sequence 1800, Ap
37	315.5	10.5	292	4	US-09-313-294A-7157	Sequence 1800, Ap
38	189.5	6.3	1374	4	US-08-415-658-70	Sequence 1800, Ap
39	175.5	5.9	2841	4	US-09-949-016-4735	Sequence 1800, Ap
40	175.5	5.9	2842	4	US-09-702-705-1800	Sequence 1800, Ap
41	175.5	5.9	2842	4	US-09-736-457-1800	Sequence 1800, Ap
42	175.5	5.8	1719	4	US-09-671-325-1800	Sequence 199, App
43	173.5	5.8	1719	4	US-09-949-016-199	Sequence 5632, Ap
44	173.5	5.8	4403	4	US-09-949-016-5632	Sequence 66, Appl
45	173.5	5.8	5421	4	US-09-814-915A-66	

ALIGNMENTS

RESULT 1
US-08-920-095-1
Sequence 1, Application US/08920095
Patent No. 5837240
GENERAL INFORMATION:
APPLICANT: Cynthia K. Lee et al.
TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,095
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/431,041
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06132/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

; LENGTH: 2735 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; US-08-920-095-1

Alignment Scores:

Pred. No.:	6,266-247	Length:	2735
Score:	2297.00	Matches:	416
Percent Similarity:	86.77%	Conservative:	76
Best Local Similarity:	73.37%	Mismatches:	75
Query Match:	76.59%	Indels:	0
DB:	2	Gaps:	0

US-09-904-994B-3 (1-568) x US-08-920-095-1 (1-2735)

```

QY      2  LysMetLysLysGlnGlnTyrValAenThrTyrGlyProThrLysGlyValAspLysValArg 21
DB      1668 GCTCTGATATATTAATAAGTAGCCGCTGAACACAACTATCTCCGCTTCCACTAACCC 1727
QY      22  LeuGlyAspThrAspLeuThrAlaGlyValGlnHisAspTyrThrThrTyrGlyGlnGlu 41
DB      1968 TCTCAAGCGATGGCGCGGTGGTGAAGTATCACTAGAACTTGGCCAAACACTGACAA 1967
QY      888  TTGGCGGATACAGACTTGATCGCTGAAGTAAGACATGACTACACCTTATGCGCGAAG 947
DB      1968 AACAGAAAGANTTTGGCCGCTTGAAGAAAGAAAGCGATACGACTTACAGACTTACAG 2027
QY      42  LeuLysPheGlyAlaGlyLysThrIleArgGlnGlyMetGlyGlnSerAsnSerProAsp 61
DB      2088 GTAGGTTCAAGTAAGAGTGGGCAAGTGCCTGGTAATGTGAGATCCACACTTCTT 2147
QY      948  CTTAATTCGTCGTGGCGGTAAACCCCTAAGAGAGCATGAGCCAACTTAACAACCTTAC 1007
DB      2088 AAACCTACTTCTTAATTAACCACTTAACCCAGCGATGATGATGATGATGATGATGAT 2087
QY      62  GluAsnThrLeuAspLeuValIleThrAsnAlaMetCileIleAspTyrThrGlyIleTyr 81
DB      1008 AAAGAAAGACTGATTAATTAATTAATCACTTAACGCTTAACTGATTAACCGGATATT 1067
QY      82  LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyValAlaGlnLys 101
DB      1068 AAAGCGGATATGATTAAGATGCAAAATCGCTGCAATGTGTAAAGCGGTAAACAA 1127
QY      102  AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGlnAlaLeuAla 121
DB      1128 GACATGCAAGATGGCGCTTAAACAACTTACGATAGCTCCGCTACTGAAAGCCTTAAG 1187
QY      122  GlyGlnGlyMetCileIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141
DB      1188 GGTGAAGGTTTATGCTTAACGCGCTGGTATTTGACACACATCCACTTTCACCC 1247
QY      142  GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161
DB      1248 CAACAAATCCCTTAAGCTTTTGCACACGGGTGAACACATGATGTGTGGAACCGGT 1307
QY      162  ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysThrAsnLeuHisArgMet 181
DB      1308 CCTGCGATGCGCATATGCGACTACTATCACTCCAGCGAAGAAATTTAAATGGAAG 1367
QY      182  LeuAlaGlnAlaGlnGlnTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB      1368 CTCAGAGCGCGTGAAGATATTTCTATGAATTTAGGTTTCTTGCGCTAAAGTAACGCTT 1427
QY      202  SerLysLysGlnLeuValGlnGlnValGlnAlaGlyAlaIleGlyPheLysLeuHisGln 221
DB      1428 AACGATGCGAGCTTGACCATTAATTTGAAGCGCGGTGACATTTGCTTAAATTCAGAA 1487
QY      222  AspTProGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241
DB      1488 GACTGGGGGACCACTCTTCTGCAATCAATCAATGCTTGAATGTTCCGGAACAAATTCGAT 1547
QY      242  ValGlnValCysIleHisThrAspThrValAsnGlnAlaGlyTyrValAspAspThrLeu 261
DB      1548 GTGCAAGTGTGATTCACACAGACACTTGAATGAAGCGGTTGTGAAGACACTATG 1607
QY      262  AsnAlaMetAsnGlyArgAlaIleHisAlaIleTyrHisIleGlnGlyValAlaGlyIleHis 281
DB      1608 GCTGCTATTTGAGACGACATGACACTTTCACACTGAAGGCGCTGCGCGGACAC 1667
  
```

```

QY      282  SerProAspValIleThrMetAlaGlyGlnLeuAsnIleLeuProSerSerThrThrPro 301
DB      1668 GCTCTGATATATTAATAAGTAGCCGCTGAACACAACTATCTCCGCTTCCACTAACCC 1727
QY      302  ThrIleProTyrThrIleAsnThrValAlaGlnHisLeuAspMetLeuMetThrCysHis 321
DB      1728 ACCATCCCTTTCACCGTGAATACAGAACGACGACATGACATGATCTTAATGCTGTGCCAC 1787
QY      322  HisLeuAspLysArgIleArgGlnAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB      1788 CACTTGGATTAAGATTAAGATTAAGATGTCATGCTTCCGCTGATTCAGAGATCCGCCCTCA 1847
QY      342  SerIleAlaIleGlnAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB      1848 ACCATTTGGCGGTGAAGACACTTTCGATGACATGGGGAATTTCTCATATCCAGTTCTGAC 1907
QY      362  SerGlnAlaMetGlyValArgAlaGlyGlnValIleProArgThrThrThrAlaPheLys 381
DB      1908 TCTCAAGCGATGGCGCGGTGGTGAAGTATCACTAGAACTTGGCCAAACACTGACAA 1967
QY      382  AsnLysLysGlnPheGlyLysLeuProGlnAspGlyLysAspAsnAspAsnPheArgIle 401
DB      1968 AACAGAAAGANTTTGGCCGCTTGAAGAAAGAAAGCGATACGACACTTACAGATC 2027
QY      402  LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlyTyr 421
DB      2088 AAACCTACTTCTTAATTAACCACTTAACCCAGCGATGCTGATGATGATGATGATGAT 2087
QY      422  IleGlySerValGlnGlnGlyLysIleAlaAspLeuValIleThrAsnProAlaPhePhe 441
DB      2088 GTAGGTTCAAGTAAGAGTGGGCAAGTGCCTGGTAATGTGAGATCCACACTTCTT 2147
QY      442  GlyValLysProLysIleValIleLysGlyLysMetValAlaPheSerGlnLysMetGlyAsp 461
DB      2148 GCGGTGAAGCCCAACATGATATCAAGAGCGGATTCATTCGCTTAAGCCAAATGGCGCAT 2207
QY      462  SerAsnAlaSerValProThrProGlnProValTyrTyrArgGlnMetPheGlyHisHis 481
DB      2208 GCGAAGCTTCTATCACTTAACCCCAACACCGGTTTATTAAGAGAAATGTTCCGCTCAT 2267
QY      482  GlyValAlaLysPheAspThrSerIleThrPheValSerLysValAlaIleTyrGlnAsnGly 501
DB      2268 GGTAAAGCTTAATTAAGATGCAAAACATCACTTGTGTCTCAAGCGGCTTATATCAAGGC 2327
QY      502  ValLysGlnLysLeuGlyLysGlnValIleProValIleAsnCysArgAsnIle 521
DB      2328 ATTAAGAGAAATTTGACACTTGAAGACAGATGTTCCGCTTAAGAAATTTGCAAGAAATATC 2387
QY      522  ThrLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
DB      2388 ACTAATAAAGACATGCAATTTCAACGACACTACTGCTCATTTGAAGTCAATCTGAAACT 2447
QY      542  PheGlnValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
DB      2448 TACCATGTGTTCTGTGAATGGCAAGAGTAATCTTAAACCGCAATTAAGTGAAGCTTG 2507
QY      562  AlaGlnArgTyrThrPhePhe 568
DB      2508 GCGCAACTCTTGAAGCAATTTTC 2528
  
```

RESULT 2

PCT-US96-05800-1
 ; Sequence 1, Application PC/TUS9605800

GENERAL INFORMATION:

; APPLICANT: Ocravex, Inc.
 ; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT URBASE VACCINE
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA

ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/05800
 FILING DATE: 23-APR-1996
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/431,041
 FILING DATE: 28-APR-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/568,122
 FILING DATE: 06-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 06132/020001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 INFORMATION FOR SEO ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2735 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 PCT-US96-05800-1

Alignment Scores:
 Pred. No.: 6,26e-247 Length: 2735
 Score: 2297.00 Matches: 416
 Percent Similarity: 86.77% Conservative: 76
 Best Local Similarity: 73.37% Mismatches: 75
 Query Match: 76.59% Indels: 0
 Gaps: 0

US-09-904-994B-3 (1-568) x PCT-US96-05800-1 (1-2735)

QY 2 Lymetlysbysgngluyrvalasnthryrtyrlyrprothryrlyasrlyvalarg 21
 Db 828 AAGATTGACAGAAAGATATGTTCTATGATGCTCTACTACAGCGATMAAGTGA 887
 QY 22 Leuglyasrthraspleutrpalaigluvalguhihsapryrthrtlyrtyrlyglu 41
 Db 888 TTGGGCGATACAGACTTGATCGCTGAAGTAGAACATGACTACCATTTATGGGAG 947
 QY 42 Leulysbheglualaglylysthrilearggluylwecgljynserasrserproasp 61
 Db 948 CTTAAATTCGGTGGCGGTAAACCTTAAGAGACATGAGCCATCTAACACCTTAC 1007
 QY 62 Gluasnthrleuasrpleuvalilethrasnalametileleasrtyrthrtlyrtyr 81
 Db 1008 AAGAGAGAGTGGATTATTAATTCATCACTAACCGTTTATCGGATTTACACCGGTATTTAT 1067
 QY 82 Lysalalasrplleglylylelyasnglylyseihiegllyilegyllyvalaglyasrly 101
 Db 1068 AAGCGGATTTGTTATTAAGATGGCAAAATTCCTGGCATTTGTAAGGCGGTACAA 1127
 QY 102 Asrmetcglasrpllyasrserprohiametvalaglyvalaglythrtgualaleu 121
 Db 1128 GACATGCAAGTGGCGTTAAACCAATCTTAGCTAGCTCTGTCTAGCAAGCTTAC 1187
 QY 122 Glygljylweclylelethrplaglylyleasrserthsrthrasrpleu 141
 Db 1188 GGTGAAGTGTGATGACGCGCTGTGTATTAACACACACATCCATTCATTTTCA 1247
 QY 142 Glnglnpheprothralaleualasnglyvalthrthrtmetphegllygllythrgly 161

Db 1248 CAACAAATCCCTACAGCTTTTGCAGAGCGGTAAACACATGATGTTGGTGAACCGGT 1307
 QY 162 Provalasrpllyrthrasnalametilethrtlyrtyrlyrprothryrlyasrly 181
 Db 1308 CTTCTATGCGCTTAATGACATCTACATCACTCAGCAGAAATTTTAAATGATG 1367
 QY 182 Leuarvalalagluglylyrsermetasrvalaglypheleuglylyasrser 201
 Db 1368 CTCAGAGCGCTTAAGATTTCTATGAATTTAGGTTCTTGGCTTAAGGTAAGCTTCT 1427
 QY 202 Serlysbysgnglueuvalagluvalaglyalaglylypelyleuhsieglu 221
 Db 1428 AAGATGCGAGCTTAGCGCATCAAAATGAGCGGTGCGATTTGCTTTAAATTCAGCA 1487
 QY 222 Asrtrpelyrthrtproseralileasrphieylyseuervalliasrpllyr 241
 Db 1488 GACTGGGGCACACCTCTTCTGCAATCAATCATCTAGATGTTGGGCAAAATACGAT 1547
 QY 242 Valgluvalaglylylethsrthrasrthrasngluvalaglylyrvalasrpllyr 261
 Db 1548 GTGCAAGTCGCTATCCACAGACACTTGAATGACCGGTGTGTAGAAAGACATATG 1607
 QY 262 Asnalamecbsnglyarvalalehialatyrthrasrlyleuglylylyly 281
 Db 1608 GCTGCTATTCGTGACACCATATGACACATTTCCACATGAGCGCTGGCGGACAC 1667
 QY 282 Serproasrvalilethrtmetalaglylyleuhsanilleuoproser 301
 Db 1668 GCTCCTATATTTATTAAGTGGCGGTGAACACAACTTTCTCCGCTCCACATACCC 1727
 QY 302 Thrileprothrtlyrthrasnthrtvalalagluhihsleuapmetleu 321
 Db 1728 ACCATCCCTTACCGGAATACAGAACAGACATGACATGCTTATGCTGTGCAC 1787
 QY 322 Hiseuasrlysbysarglilearggluasrpleuaglinserserarglilearg 341
 Db 1788 CACTTGATTAAGCATTAAGAAAGATGTTCACTTCCCTGATTCAGAGATCCGCTCAA 1847
 QY 342 Serilealalagluasrvalileuhsasrmetcgllyalilealmet 361
 Db 1848 ACCATTCGCGCTGAAGACACTTTCATGATGAGGAGATTTCTCAATCACCATGCTGAC 1907
 QY 362 Sercglinalmetcgllyarvalaglylyvalileproarghrttrp 381
 Db 1908 TCTCAAGCGATGGCGCTGTGGGTGAATCATCTGAACCTGGCAACGCTGACAA 1967
 QY 382 Asnlysbysgngluphegllylyleuoprogluasrpllyasrpaasrpaasr 401
 Db 1968 AACAGAAAGATTTGGCGCTTGAAGAAAGAAAGCGATTAACGAACTTCAGATC 2027
 QY 402 Lysargtyrlyleserlyrthrtleasrproalaleuthrthrasrly 421
 Db 2028 AAAGCTACTTGTCTAATATACACCATTAACCGAGATTCCTGAGGATTAAGGAT 2087
 QY 422 Illeglyservalagluglylylylealasrpleuvaliltpasrproalapher 441
 Db 2088 GTAGTTCAGTAGAAGTGGCAAGTGGCTGACTGTGATTTGGAATCCAGCATTTCTT 2147
 QY 442 Glylylysbysrpllylevalilelylegllylymetvalalphe 461
 Db 2148 GCGGTGAACCCACATGATCATCAAGCGGATTCCTGCTTAAGCCAAATGGCGAT 2207
 QY 462 Serasnalaservalprothrtproglinservallyrtyrarglumetpheglly 481
 Db 2208 GCGAACCTTCTATCCCTACCCCAAGCGTTTATTAAGAGAAATGTTGCTCATCAT 2267
 QY 482 Glylyvalasrpleasrpllyrthrtphervalaserlyvalalalyr 501
 Db 2268 GGTAAAGCTTAATACGATGCAACATCATCTTGTGTCTCAAGCGGTATGACAAAGC 2327
 QY 502 Vallysglylysbysgnglyleugluarggluvalleuoprovallyasr 521

Db 2338 ATTAAGAGAAATTAAGACTTGAAGACAAAGTGTCCGGTAATAAAATTGCAGAAATATC 2387
 QY 522 ThrLVsAspPheLVsPheAsnAspLVsThrAlaLVsIleThrValAspProLVsThr 541
 Db 2388 ACTAATAAAAGACATGCAATTCACGACACTGCTCACAATGAGATCAATCTGAAACT 2447
 QY 542 PheGluValPheValAspGlyLVsLeuCy8ThrSerLVsProThrSerGluValProLeu 561
 Db 2448 TACCATGTGCTGTGGATGGCAAGAAAGTAACCTTCAACCAAGCCAAATAAGTAGACTTG 2507
 QY 562 AlaGluArgTyrThrPhePhe 568
 Db 2508 GCGCAACTCTTACGATTTTC 2528

RESULT 3

US-09-431-705-1

/ Sequence 1, Application US/09431705
 / Patent No. 6585975
 / GENERAL INFORMATION:
 / APPLICANT: Kleantinous, Harold
 / APPLICANT: Londono-Arcila, Patricia
 / TITLE OF INVENTION: Use of salmonella vectors for
 / FILE REFERENCE: 06132/060001
 / CURRENT APPLICATION NUMBER: US/09/431,705
 / NUMBER OF SEQ ID NOS: 52
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 4824
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: includes sequences from *Helicobacter pylori*,
 / OTHER INFORMATION: *Salmonella typhimurium*, and *Escherichia coli*
 / NAME/KEY: CDS
 / LOCATION: (2)...(31)
 / NAME/KEY: CDS
 / LOCATION: (41)...(61)
 / NAME/KEY: CDS
 / LOCATION: (65)...(799)
 / NAME/KEY: CDS
 / LOCATION: (803)...(2512)
 / NAME/KEY: CDS
 / LOCATION: (2516)...(2692)
 / NAME/KEY: CDS
 / LOCATION: (2696)...(2896)
 / NAME/KEY: CDS
 / LOCATION: (2900)...(3322)
 / NAME/KEY: CDS
 / LOCATION: (3326)...(3385)
 / NAME/KEY: CDS
 / LOCATION: (3389)...(3406)
 / NAME/KEY: CDS
 / LOCATION: (3410)...(3466)
 / NAME/KEY: CDS
 / LOCATION: (3470)...(3598)
 / NAME/KEY: CDS
 / LOCATION: (3602)...(3661)
 / NAME/KEY: CDS
 / LOCATION: (3665)...(3697)
 / NAME/KEY: CDS
 / LOCATION: (3701)...(3769)
 / NAME/KEY: CDS
 / LOCATION: (3773)...(3817)
 / NAME/KEY: CDS
 / LOCATION: (3821)...(3844)
 / NAME/KEY: CDS
 / LOCATION: (3848)...(3889)
 / US-09-431-705-1

Alignment Scores:

Pred. No.: 9,78e-245 Length: 4824
 Score: 2281.00 Matches: 414
 Percent Similarity: 86.42% Conservative: 76
 Best Local Similarity: 73.02% Mismatches: 77
 Query Match: 76.06% Indels: 0
 DB: 4 Gaps: 0

US-09-904-994b-3 (1-568) x US-09-431-705-1 (1-4824)

QY 2 LysMetClyLVsGluGluTyrValAsnThrTyrGlyProThrLVsAspLVsValArg 21
 Db 812 AAGATTAGCAGAAAGAAATATGTTCTATGATGGTCTTACTACAGCGCATTAAGTGACA 871
 QY 22 LeuGlyAspThrAspLeuTyrValGluValGluIleAspTyrThrTyrGlyGluGlu 41
 Db 872 TTGGGCGATACAGACTTGATCGCTGAGTAGAACAATGACTACACCATTTATGCGCAAGG 931
 QY 42 LeuLVsPheGlyValGlyLVsThrIleArgGlyMetGlyGlnSerAsnSerProAsp 61
 Db 932 CTTAAATTCGTGGCGGCTTAAACCTTAAGAGAGAGCAGACCAATCTAACCAACCTTAC 991
 QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleAspTyrThrGlyIleTyr 81
 Db 992 AAGAGAGATTGATTTAATTAATCACTAACGCTTAAATGCTGATTTACACCGGATTTAT 1051
 QY 82 LysAlaAspIleGlyIleLVsAsnGlyLVsIleHisGlyIleGlyLVsAlaGlyAsnLVs 101
 Db 1053 AAGCGGATATTGATTTAATTAATGAGCAAAATCGCTGGATTTGTAAGCGGTAACAA 1111
 QY 102 AspMetGluAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeuAla 121
 Db 1112 GACATGCAAGATGGCGGTTAAAAACAATCTTACCGTCTGATCGTCACTGAAACCTTAGCC 1171
 QY 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141
 Db 1172 GGTGAAGGTTGATGTAACGCGCTGTGTGATTTGACACACATCCACTTCATTCACCC 1231
 QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161
 Db 1232 CAACAAATCCCTACAGCTTTTGCAAGCGGTGACACATGATTTGGTGGAAACCGGT 1291
 QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLVsTyrAsnLeuHisArgMet 181
 Db 1292 CCTGCTGATGGCATAATGCGACTACTACCTCCAGGCAAGAAATTTAAATGGAAG 1351
 QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLVsGlyAsnSerSer 201
 Db 1352 CTCAGAGCGGCTGAAGAAATTTCTATGAATTAAGTTCTTGGCTTAAGTAACGCTTCT 1411
 QY 202 SerLVsLVsGluLeuValGluGluValGluAlaGlyAlaIleGlyPheLeuHisGlu 221
 Db 1412 AACGATGCGAGCTTAGCGCATGAATTAAGCGGATTTGGCTTGCATTTACAGAA 1471
 QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241
 Db 1472 GACTGGGGCACCACTCTTCTCAATCAATCAATGAGTTGAGATTTGGCGGCAAAATACGAT 1531
 QY 242 ValGluValCysIleHisThrAspThrValAsnGluValGlyTyrValAspAspThrLeu 261
 Db 1532 GTGCAAGTCCTATGCGCACACACACTTGAATGAAGACCGGTTGGTGAACACACTATG 1591
 QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValGlyGlyGlyHis 281
 Db 1592 GGTGCTATTGCTGACGCACTATGACACTTTCCACATGAAGGCGCTGGCGGCGACAC 1651
 QY 282 SerProAspValIleThrMetAlaGlyLeuAsnIleLeuProSerSerThrThrPro 301
 Db 1652 GCTCTGATATTAATTAAGTAGCCGATGAACACACATCTTCCGCTTCCACTAACCCC 1711
 QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisIleAspMetLeuMetThrCysHis 321
 Db 1712 ACCATCCCTTTTACCGTGATATCAGAAAGACAGACATGACATGATGATGATGTCACAC 1771

QY	DB	Sequence	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	Db	1052 AAGCGGATATGGTATTAAGATGGCAAAATCGCTGCATTTGGTAAGCGGTAAACAA	1111	4824	414	76	77	0	0
QY	Db	102 ASPMeGlnASPGLVValSerProHISmetValValGlyValGlyThrGlnAlaLeuAla	121	4824	414	76	77	0	0
QY	Db	1112 GATTCAGAGATGGCGTTAAAAACAATCTTACGTAAGTCTCTGTAAGCCTTAGCC	117	4824	414	76	77	0	0
QY	Db	122 GYGLGlyMetIleIleThrAlaGlyGlyIleAspSerHISThrHISpHeLeuSerPro	141	4824	414	76	77	0	0
QY	Db	1172 GGTGAAGTGTATGATCGTAACGGCTGTGTGTAACAACATCCATTCATTTCAACC	123	4824	414	76	77	0	0
QY	Db	142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyIleThrGly	161	4824	414	76	77	0	0
QY	Db	1232 CAACCAATCCCTACAGCTTTTGCAGGCGGTGTAAACAACATGATGTGTGTGAACCGGT	129	4824	414	76	77	0	0
QY	Db	162 ProValASPGLVThrAsnAlaThrThrIleThrProGlyIleValThrPheLeuHISArgMet	181	4824	414	76	77	0	0
QY	Db	1292 CTCGTCGATGAGCACTAAATCCGATCTATCATCTCCAGGCAAGAAATTTTAAATGATG	135	4824	414	76	77	0	0
QY	Db	182 LeuArgAlaIleGlyGlyIleValSerMetAsnValGlyPheLeuGlyIleGlyAsnSerSer	201	4824	414	76	77	0	0

Db 1352 CTCAGACGGCTGAGAAATATCTATGAATTAGGTTCTTGCGTAAGGTAAAGCTTCT 1411
 QY SerTylsYlsgInleuValgluInValaglValaileglYpHeuLsleuHlsglu 221
 Db 1412 AACGATGCGAGCTTACCGCATCAATTTGAAGCCGGTGGCATTTGGCTTGCATTCAGAA 1471
 QY 222 AapTlrglYlThrThrProSerAlaileaspHisCylseuSerValAlaaspGluYraap 241
 Db 1472 GACTGGGGACCACTCTTCTGCAATCAATCATGCGCTTGAATGTTGCGCAAAATCGAT 1531
 QY 242 ValgluValCylleHlshThrAspThrValAsnGluAlaglYlThrAlaaspThrLeu 261
 Db 1532 GTGCAAGTGGCTATCGCCACAGACCTTGAATGAAGCCGGTGTGTGAAGAGACTATG 1591
 QY 262 AsnAlaMetAsnGlyAArgAlaileHlslalAryHsllegluGlyAlaglYlglYHls 281
 Db 1592 GCGGCTATTGCTGAGCACTATGCACACTTTCACACAGGAAGCGCTGGCGGCGGACAC 1651
 QY 282 SerProAspValilleThrMetAlaglYlgluLeuAsnilleuProSerSerThrPro 301
 Db 1652 GCTCCGATATATATTAAGTAGCCGGTGAACACAAATCTTCCCGCTTCACTAACCCC 1711
 QY 302 ThrIleProDylThrIleAsnThrValAlaglHlsleuAspMetLeuMetThrCysHls 321
 Db 1712 ACCATCCCTTTCACCGGTAAATACAGAGACAGACATGACATGCTTATGTGTGCCAC 1771
 QY 322 HlsleuAspYlsgAArgIleAArgIleuAspLeuGlnPheSerGlnSerAArgIleAArgProGly 341
 Db 1772 CACTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1831
 QY 342 SerIleAlaAlaglAspValleuHlsaspMetGlyValilleAlaMetThrSerSerAap 361
 Db 1832 ACCATGCGCGGTGAACACTTTCGATGATGAGGAGATTTTTCATCATCACAGTTCGAC 1891
 QY 362 SerGlnAlaMetGlyAArgIleAArgIleuValilleProAArgThrTrpGlnThrAlaAspYls 381
 Db 1892 TCTCAAGCCATGGCGCGGTGGTGAAGTATCATGAACTTGGCAACAGCTGACAA 1951
 QY 382 AsnYlsgYlsgInleuPheGlyLysleuProGluAspGlyLysAspAsnAspAsnAspAsp 401
 Db 1952 AACAAAGAAATTTGGCCGCTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2011
 QY 402 LysAArgTylleSerYlsgYlThrIleAsnProAlaLeuThrHlsglyValSerGluYr 421
 Db 2012 AAACGCTACTTGTCTAAATACACATTAACCCGACATGCTCATGAGATTAAGCAGATAT 2071
 QY 422 IleglYserValgluInleuYlsgIleAaspLeuValValTrpAsnProAlaPhePhe 441
 Db 2072 GTAGGTTCAAGTGAAGTGGCAAGGCTGAGCTGTGATGTGGAGTCCAGCATTTCTT 2131
 QY 442 GlyValYlsgProLysIleValilleYlsgYlglYMetValValPheSerGluMetGlyAap 461
 Db 2132 GCGGTAAACCCCAACATGATCATCAAGCGGATTCATTCGCTTAAGCCAAATGGCGCAT 2191
 QY 462 SerAsnAlaSerValProThrProGlnProValYlsgYlglYMetPheglYHlsHls 481
 Db 2192 GCGAAGCTTCTATCCCTACCCCAACGCGTTTATTATTAAGAAATGTTCGCTCATAT 2251
 QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerYlsgValAlaYlrgluAsnGly 501
 Db 2252 GGTAAAGCTAAATACATGCAACATCATCTTTGTCTCAACGCGCTTAATGCAAAAGGC 2311
 QY 502 ValYlsgYlsgLeuGlyLeuGluAArgGlnValleuProValYlsgAsnCysAArgAsnIle 521
 Db 2312 ATTAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2371
 QY 522 ThrYlsgYlsgAspPheAspThrPheAsnAspYlsgThrAlaLysIleThrValAlaaspProLys 541
 Db 2372 ACTAAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2431
 QY 542 PheGluValAlaPheValAspGlyLysleuCysThrSerYlsgProThrSerGlnValProLeu 561
 Db 2432 TACCATGTGTTCGTGATGGCAAGAAAGTAACTTTTAAACCAACCAATAAAGTGAAGCTTG 2491

QY 562 AlaGlnAArgYlThrPhePhe 568
 Db 2492 GCGCAACTCTTATGACATTTTC 2512
 RESULT 5
 US-08-467-822-19
 ? Sequence 19, Application US/08467822
 ? Patent No. 5843460
 ? GENERAL INFORMATION:
 ? APPLICANT: Labigne, Agnes
 ? APPLICANT: Sauerbaum, Sebastien
 ? APPLICANT: Ferrero, Richard L.
 ? APPLICANT: Thiberge, Jean-Michel
 ? TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 ? TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 ? TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 ? TITLE OF INVENTION: POLYPEPTIDES
 ? NUMBER OF SEQUENCES: 44
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ? ADDRESSEE: Dunner
 ? STREET: 1300 I Street, N.W.
 ? CITY: Washington
 ? STATE: D.C.
 ? COUNTRY: USA
 ? ZIP: 20005-3315
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patentin Release #1.0, Version #1.30
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/467,822
 ? FILING DATE: 06-JUN-1995
 ? CLASSIFICATION: 435
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/447,177
 ? FILING DATE: 19-MAY-1995
 ? CLASSIFICATION: 435
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/432,697
 ? FILING DATE: 02-MAY-1995
 ? CLASSIFICATION: 435
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Meyers, Kenneth J.
 ? REGISTRATION NUMBER: 25,146
 ? REFERENCE/DOCKET NUMBER: 03495.0137-02000
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (202) 408-4000
 ? TELEFAX: (202) 408-4400
 ? INFORMATION FOR SEQ ID NO: 19:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 2619 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: double
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: DNA (genomic)
 ? FEATURE:
 ? NAME/KEY: misc_feature
 ? LOCATION: 31..36
 ? OTHER INFORMATION: /standard_name= "Shine-Dalgarno
 ? FEATURE:
 ? NAME/KEY: misc_feature
 ? LOCATION: 756..759
 ? OTHER INFORMATION: /standard_name= "Shine-Dalgarno
 ? OTHER INFORMATION: sequence."

Alignment Scores: 5.14e-241 Length: 2619
 Pred. No.: 2244.00 Matches: 411
 Score:

Percent Similarity: 85.01%
 Best Local Similarity: 72.49%
 Query Match: 74.82%
 DB: 2
 Gaps: 0

US-09-904-994B-3 (1-568) x US-08-467-822-19 (1-2619)

Conservative: 71
 Mismatches: 85
 Indels: 0

QY 2 LysMetLysbSGInGLuTYrValAsnThrTYrGlyProThrLysGlyAspLysValArg 21
 Db 772 AAGATTTCACGAAAAAGATATGTTTCTATGTATGTGCTCCACTACCGGGAGATGTTTGA 831
 QY 22 LeuGlyAspThrAspLeuTPDAlaGluValGlnHisAspTYrThrTYrGlyGlu 41
 Db 832 CTCGGCGACACTGATTTGATCTTGAAGTGAGCATGATGACACCACTTAAGTGAAG 891
 QY 42 LeuLysPheGlyAlaGlyLysThrLLeaRgGluGlyMetGlyGlnSerAsnSerProAsp 61
 Db 892 ATCAAAATTTGGGGCGGTAAACATATCCGTATGGATGGATCAAAACCAATATGCCCTAGC 951
 QY 62 GluAsnThrLeuAspLeuValLLeaThrAsnAlaMetLLeaAspTYrThrGlyLLeaTYr 81
 Db 952 TCTTATGAAATTAATTTGCTGCTCACTAACCCCTCATTTGTGACTATACGGCATTTTAC 1011
 QY 82 LysAlaAspLLeaGlyLLeaLysAsnGlyLysLLeaGlyLLeaGlyLysAlaGlyAsnLys 101
 Db 1012 AAAGCCGACATTTGGGATTAAGACGCAAGATTTGACGACATTTGGCAAGGCAATAG 1071
 QY 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeuAla 121
 Db 1072 GACATGCAAGATGCGTATGATATATCTTTGCTGAGTCCGTGCTACAGAGGCTTTGCA 1131
 QY 122 GlyLysGlyMetLLeaLLeaThrAlaGlyLLeaAspSerHisThrHisPheLeuSerPro 141
 Db 1132 GCTAGAGGCTTGAATTTGATGACCGGTGGTGCATGCAATGCAATTCATCTTATCTCTCC 1191
 QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThrGly 161
 Db 1192 CAACAAATCCCTATCGCTTTGCCAGCGGGTTTCAACATGATTTGGAGGAGGCAAGCA 1251
 QY 162 ProValAspGlyLLeaAsnAlaThrThrLLeaProGlyLysTPRAsnLLeaHisArgMet 181
 Db 1252 CCTCGGATGGCAGAAATGCAACCACTCATCTCCGAGCGCTTAATCTTAAAAAGTATG 1311
 QY 182 LeuArgAlaAlaGluGluTYrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
 Db 1312 TTGGGTGAGCGCAAGATACGCAATGATCTAGCTTTTGGCTTAAGGGGAATGTGCT 1371
 QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaGlyPheLysLeuHisGlu 221
 Db 1372 TAGCAACCCCTTTACGGCATGATTCAGACAGCGGGCATTTGTTTAAAAATCCAGCA 1431
 QY 222 AspTyrGlyLLeaThrProSerAlaLLeaAspHisCysLeuSerValAlaAspGluTYrAsp 241
 Db 1432 GACTGGGGAGACACCTCGACGATTTACCACTGCTCATATGTCGCGAGTAATACGAT 1491
 QY 242 ValGlnValCysLLeaLLeaThrAspThrValAsnGluAlaGlyTYrValAspAspThrLeu 261
 Db 1492 GTGCAAGGCTTATCCACACCGATACCTTAAACAGAGGGGCTGTGTGAAGAACCCCTA 1551
 QY 262 AsnAlaMetAsnGlyArgAlaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLea 281
 Db 1552 GAGCGCATTTGGCGGGCGGACATCATCTTCCACACTGAAGGGGCTGGGGGTGAGCAC 1611
 QY 282 SerProAspValLLeaMetAlaGlyLLeuLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLea 301
 Db 1612 GCTCCATGATTTATCAAAATTTGGCAGGGGAATTTAACTTACCCGCTTACTTAACCG 1671
 QY 302 ThrLLeaProTYrThrLLeaThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
 Db 1672 ACCATTCCTTTCACCAAAACACGAAACCGACGACATGACATGTAATGATGTGGCAC 1731
 QY 322 HisLeuAspLysArgLLeaRgGluAspLeuGlnPheSerGlnSerArgLLeaRgProGly 341

Db 1732 CACTTGGATAAAGATATCAAGAAAGATGTCAGTTTGGCCATTCGAGGATTCGCCCCAA 1791
 QY 342 SerLLeaAlaAlaGluAspValLLeuHisAspMetGlyValLLeaLLeaMetThrSerAsp 361
 Db 1792 ACTATCCGGGCTGAAGCAACCTCCATGACATGGGAGATCTTTTCTATCCACAGCTCCGAC 1851
 QY 362 SerGlnAlaMetGlyArgAlaGlyLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLea 381
 Db 1852 TCTAGGCTATGGAGACCGCTAGGCGAGATGATCAACGCACTTGGCAGACAGCAACAA 1911
 QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgLLe 401
 Db 1912 AACAAAAAGATTTGGCGGCTTGAAGAGAAAAAGCCGATTAACGCAACTTCGCGATC 1971
 QY 402 LysArgTYrLLeaSerLysTYrThrLLeaAsnProAlaLeuThrHisGlyValSerGluTYr 421
 Db 1972 AAAGCTACATCTTAAATACACATCAACCCCGGATCCGCGATGGGATTTGCACTAT 2031
 QY 422 LLeaGlySerValGluGluGlyLysLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLea 441
 Db 2032 GTGGGCTCTGAGAAAGTGGCAATACGCGACCTCGTGTGAGATCCGGCTTTCTT 2091
 QY 442 GlyValLysProLysLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLeaLLea 461
 Db 2092 GGCAATTAAGCCCATATGATTAATTAAGGCGGATTTATGCGCTCTCAAAATGGCGAT 2151
 QY 462 SerAsnAlaSerValProThrProGlnProValTYrTYrArgGluMetPheGlyHis 481
 Db 2152 GCCAATGCTTATTCACACCCCTCACCCGCTCTATTAACCGTGAAGTGTGGACACAT 2211
 QY 482 GlyLysAlaLysPheAspThrSerLLeaThrPheValSerLysValAlaTYrGluAsnGly 501
 Db 2212 GGGAAAAACAAATTCGACCAATATCACTTCTGTGTCGCAAGGGCTTTACAGGCAAGG 2271
 QY 502 ValLysGluLysLeuGlyLLeuGluArgGlnValLeuProValLysAsnCysArgAsnLLe 521
 Db 2272 ATCAAGAAAGAACTAGAGGCTATGATCGCGGCAACGCGCAAGTGAAGAACTGTGCAATATC 2331
 QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysLLeaLLeaLLeaLLeaLLea 541
 Db 2332 ACTAAAAAGGCTCAAAATTCAGAGATGACCCGACATTTGTATGTCACACCTGAAACC 2391
 QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
 Db 2392 TATAGGTGAAGATGATGGCAAGAGATTACTTTAAAGCAGATGAATTAAGCTTA 2451
 QY 562 AlaGlnArgTYrThrPhePhe 568
 Db 2452 GCGCACTTATATTTGTTT 2472

RESULT 6
 US-08-432-697-19
 ; Sequence 19, Application US/08432697
 ; Patent No. 6248330
 GENERAL INFORMATION:
 APPLICANT: Labigne, Agnes
 APPLICANT: Sauerbaum, Sebastian
 APPLICANT: Ferreira, Richard L.
 APPLICANT: Thibierge, Jean-Michel
 TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 NUMBER OF SEQUENCES: 44
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
 ADDRESS: Dunnet
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..36
OTHER INFORMATION: /standard_name="Shine-Dalgarno"
OTHER INFORMATION: sequence."
NAME/KEY: misc_feature
LOCATION: 756..759
OTHER INFORMATION: /standard_name="Shine-Dalgarno"
OTHER INFORMATION: sequence."
US-08-432-697-19

```

```

Alignment Scores:
Pred. No.: 5.14e-241 Length: 2619
Score: 2244.00 Matches: 411
Percent Similarity: 85.01% Conservative: 71
Best Local Similarity: 72.49% Mismatches: 85
Query Match: 74.82% Indels: 0
DB: 3 Gaps: 0

```

US-09-904-994b-3 (1-568) x US-08-432-697-19 (1-2619)

```

QY 2 LysMetLysLeuGlnGlyValAsnThrTyrGlyProthrLysGlyAspLysValArg 21
DB 772 AAGATTTCAGAAAGAAATGTTTCATGATGTCACATCCGGGAGTCGTGAAG 831
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrThrTyrGlyGlu 41
DB 832 CTCGGGAGACCTGATTGATCTTGAAGAGTGGACATGATGGACCACTTATGCTGAAG 891
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
DB 892 ATCAAAATTTGGGGCGGTAAACATATCCGTGATGGATGAGTCAAACCAATATCCCTAGC 951
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81
DB 952 TCTTATGAATTAATATTGCTGCTACATACGCCCTCATTTGAGCATATACGGGCATTTC 1011
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
DB 1012 AAAGCGGACATTTGGATTAAAGACGGCAAGATTGGACGATTTGGCAAGGACGCAATTAAG 1071
QY 102 AspMetGlnAspGlyValSerProHisMetValValGlyValGlyLysThrGluAlaLeuAla 121
DB 1072 GACATCAAGATGCGCTAGATTAATATCTTTCGTAGTCTCTCTACAGAGGCTTTGGCA 1131
QY 122 GlyGluGlyMetIleLeuThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
DB 1132 GCTGAAGGCTTGTATTAACCGCTGGGTCATGATGATTAATCACTTATCTCTCCC 1191
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyLysThrGly 161

```

```

DB 1192 CAACAATCCCTACGCTTTTCCAGCGGGGTAAACACCATATTTGAGAGGACAGAA 1251
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
DB 1252 CCGCGGATGCGACGAATGCGACACCATCTCCCGAGCGGCTTATTTAAAGATATG 1311
QY 182 LeuArgAlaAlaGlyGluLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB 1312 TTGCGTGACGCGGAAGAAATAGCCATGATCTAGGCTTTTGGCTTAAGGGAATGTCT 1371
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGly 221
DB 1372 TAGCAACCTCTTTACGCATCAGATTGAAGACGGGGCATTTGGTTTAAATCCACGAA 1431
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLysSerValAlaAspGlyTrpAsp 241
DB 1432 GACTGGGGAAGCACCTGACACTATTCACACCTGCTCAATGTGCGCATTAATACGAT 1491
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261
DB 1492 GTGCAAGTGGCTATCCACACGATACCTTAACGAGGGCGGTGTAGAACACCCCTA 1551
QY 262 AsnAlaMetAsnGlyAspAlaIleHisAlaTyrHisIleGlyLysAlaGlyLysHis 281
DB 1552 GAGCGATTTGCCGGCGCACCATCATCTTCCACACAGAAAGGGGCTGGGGGTGACAC 1611
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrPro 301
DB 1612 GCTCCAGATGTTATCAAAATGACAGGGGAATTTAATCATTTACCGGCTTACTAACCG 1671
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLysAspMetLeuMetThrCysHis 321
DB 1672 ACCATTCTTTACCAAAAACACTGAAGCGGACATGATGATGATGATGATGATGATGAT 1731
QY 322 HisLysAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 1732 CACTGGATTAATAAGATCAAGAAAGATGAGTGGCATTTGCCGATTCGAGATTCGCCCAA 1791
QY 342 SerIleAlaIleGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 1792 ACTATCGCGGCTGAAGACCAATCCATGACATGGGATTTTCTATACCAAGCTCCGAC 1851
QY 362 SerGlnAlaMetGlyAspAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
DB 1852 TCTCAGCGCTATGAGACGCTTACGCGATGATCACACGACCTTGGAGACACAGACAA 1911
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
DB 1912 AACAAAAAAGATTTGGCGCTTGAAGAGGAAAAAGCGGATTAACGACAACTTCGCGCATC 1971
QY 402 LysArgTyrIleSerLysThrIleAsnProAlaLeuThrHisGlyValAspSerGlyTyr 421
DB 1972 AAACGCTATATCTTAATAATACCATCAACCCCGGAGTGGCATGAGGATTTCTCATAT 2031
QY 422 IleGlySerValGluGluLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
DB 2032 GTGGGCTCTGTGGAAGTGGGCAAAATACGCGCACCTCGTCCTTGGAGTCGGCTTTCTT 2091
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
DB 2092 GGCATTAAAGCCCAATATGATTAATTAAGGGCGGATTAATGACCTCTCAAAATGGCGCAT 2151
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHisHis 481
DB 2152 GCCAATGCGCTTATCCACCCCTCAGCCGCTTATACCGTGAATGTTTGGACACCAT 2211
QY 482 GlyValAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501
DB 2212 GGGAAAAAATAATTTGACACCAATATCACTTTCGTGTCCCAAGCGGCTTACAAAGGAGG 2271
QY 502 ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnIle 521

```

Db 2272 ATCAAGAGAACTAGGCTAGATCGCGGACCGCAGTGAAGAACTGTGCATATTC 2331
Qy 522 ThrlyslAspPhePheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2332 ACTTAAAGAGACCTCAATTAACGATGTGACCGACATATTGATGTCAACCTGAACCC 2391
Qy 542 PheGluValPheValAspGlyLysLeuCyThrSerLysProThrSerGluValProLeu 561
Db 2392 TATTAAGGTGAAGAGGATGCGAAGAGGTAACTCTTAAGACAGCATGAATTAAGCTTA 2451
Qy 562 AlaGluAspGlyThrPhePhe 568
Db 2452 GCGCACTTATATTAATTGTTC 2472

RESULT 7
US-08-466-248-19
Sequence 19, Application US/08466248
Patent No. 6258359
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thibierge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fimegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,248
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
APPLICATION NUMBER: 435
APPLICATION DATA:
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..36
OTHER INFORMATION: /standard_name="Shine-Dalgarno
OTHER INFORMATION: sequence."

NAME/KEY: misc_feature
LOCATION: 756..759
OTHER INFORMATION: /standard_name="Shine-Dalgarno
OTHER INFORMATION: sequence."
US-08-466-248-19

Alignment Scores:
Pred. No.: 5,14e-241 Length: 2619
Score: 2244.00 Matches: 411
Percent Similarity: 85.01% Conservative: 71
Best Local Similarity: 72.49% Mismatches: 85
Query Match: 74.82% Indels: 0
Gaps: 0
DB: 3

US-09-904-994b-3 (1-568) x US-08-466-248-19 (1-2619)

Qy 2 LysMetLysLysGlnGluLysThrValAsnThrTyGlyProThrLysGlyAspLysValArg 21
Db 772 AAGATTTCACGAAAGATATGTTTCATGTATGTGCTCCACTACCGGGGATCGTTAGA 831

Qy 22 LeuGlyAspThrAspLeuThrPalaGluValGluHisAspLysThrThrTyGlyGluGlu 41
Db 832 CTCGGCAGACACTGATTTGATCTTGAAGGTGAGCATGATTCACACTTATGTGAAGAG 891

Qy 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
Db 892 ATCAAAATTTGGGGCGGTAAACCTATCCGTGATGGATGAGTCAAACCATAGCCCTAGC 951

Qy 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspLysThrGlyIleTyR 81
Db 952 TCTATGATTAATTAAGATTGGTGCTCATTACGCCCTCATTTGGACTATACGGGATTTAC 1011

Qy 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysValAlaGlyAsnLys 101
Db 1012 AAAGCCGACATTTGGGATTTAAAGCGCAAGATTCAGCATTTGCCAAGCGCAATTAAG 1071

Qy 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyLysThrGluAlaLeuAla 121
Db 1072 GACATGCAAGATGGCGTATGATTAATATATCTTGGCTAGCTCTGCTACAGAGCTTTGGCA 1131

Qy 122 GlyGluGlyMetIleIleThrAlaGlyIleLysAspSerHisThrHisPheLeuSerPro 141
Db 1132 GCTAGGGCTTGAATTTGATTAACGGCTGGTGCATGCATGATTAATTAATCTTATCTGCC 1191

Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyLysThrGly 161
Db 1192 CAACAAATTCCTACTGCTTTTGGCCAGCGGGTTTCAACATGATTTGAGAGGACACAGA 1251

Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysThrAsnLeuHisArgMet 181
Db 1252 CTGCGGATGGCAGAAATGGAGACCATCATCTCCGAGCGCTTAATCTTAAGAAAGTATG 1311

Qy 182 LeuArgAlaAlaGluGluLysThrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
Db 1312 TTGGGTGACGCGAAGATAGCCCATGATCTAGGCTTTTGGTAAAGGGAATGTGTCT 1371

Qy 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysPheLeuHisGlu 221
Db 1372 TACGAACCTCTTATCCGCTATCAGATTTGAAGCAGGGCGATTTGGTTTAAATCCAGAA 1431

Qy 222 AspThrGlyThrThrProSerAlaIleAspHisGlyLeuSerValAlaAspGluTyRAsp 241
Db 1432 GACTGGGGAAGCACACTGACGCTATTCACACATGCTCTCAATGTCGCGAATGATAGAT 1491

Qy 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyLysThrValAspAspThrLeu 261
Db 1492 GTGCAATGGCTATCCACACGATACCTTAAGAGCGGGCTGTGTGAAGAACACCTTA 1551

Qy 262 AsnAlaMetAsnGlyAlaGlaIleHisAlaTyRHisIleGluGlyAlaGlyGlyLysHis 281
Db 1552 GAGCGATTGCGGGCCACCATCTCAATCTTCCACACTGAAGGGGCTGGGGGTGACAC 1611

Qy 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301

```

Db      1612 GCTCCAGATGATATCAAAATGCGAGGGAATTTAACTTACCCGCTCTACTAACCCG 1671
Qy      302 ThrileProtyrThrileanThrValaIaGlunHleuAspMetleuMetThrCySHis 321
Db      1672 ACCATTCCTTTCACCAAAAACACTGAAGCCGAGCAATGACATGTTAAATGGTGCCAC 1731
Qy      322 HisLeuAspLysArgIleArgGluAAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db      1732 CACTTGATTAAGTACAGAAAGATGCGCACTTGGCCGATTCGAGATTCGCCCCCAA 1791
Qy      342 SerIleAlaIleGluAAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
Db      1792 ACTATGCGCGCTGAAGACCAACCATCCATGACATGGGATCTTTCTATCAACAGCTCCGAC 1851
Qy      362 SerGlnAlaMetGlyValArgAlaGlyValIleProArgThrTrpGlnThrAlaAspLys 381
Db      1852 TCTCAGGCTATGAGGAGCGCTAGCGGATGATCAGCACTTGGCAGACAGACAGACAA 1911
Qy      382 AsnLysLysGluPheGlyLysLeuProGluAAspGlyLysAspAsnAspAspPheArgIle 401
Db      1912 AACAAAGAGAGTTGGCGCTTGAAGAGAGAAAGCGATACACAACTTCCGATC 1971
Qy      402 LysArgIleIleSerLysTrpThrIleAsnProAlaLeuThrHisGlyValSerGlnTrp 421
Db      1972 AACCGTACATCTCTAAATACACATCAACCCCGGATCGCGCATGGATTTCTGACTAT 2031
Qy      422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
Db      2032 GTGGGCTCTGTGAGAAATGGGCAAAATCGCGCACTGTGCTGTGGAGTCCGGCTTCTTT 2091
Qy      442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGlnMetGlyAsp 461
Db      2092 GGCATTACGCCCAATGATGATTTAAAGCGGATTTATTCGCTCTCAATGGCGCAT 2151
Qy      462 SerAsnAlaSerValProThrProGlnProValTrpTrpArgGluMetPheGlyHis 481
Db      2152 GCCAATGCGCTATTCGCCACCCCTCAGCCGCTGTATTAACGTAAATGTTGGACACCAT 2211
Qy      482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTrpGluAsnGly 501
Db      2212 GGGAAAAACAAATTCACACCAATATCACCTTCTGCCAACCGGCTTACAGAGCGAG 2271
Qy      502 ValLysGluLysLeuGlyLeuGluArgGlnValLeuProValLysAsnCyAspAsnIle 521
Db      2272 ATCAAGAAAGAACTAGAGCTAGATCCGCGGACCCCGCTGAAAACCTGCGCAATATC 2331
Qy      522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysTrp 541
Db      2332 ACTAAAAAGACCTCAAAATTCACACGATGACCGCACATATGATGTCACACCTGAACC 2391
Qy      542 PheGluValPheValAspGlyLysLeuCyThrSerLysProThrSerGlnValProLeu 561
Db      2392 TTAAAGTGAAGAGTGGCAAGAGGTAACTCTTAAGCAGACAGATGAATTGAGCCTA 2451
Qy      562 AlaGlnArgTrpThrPhePhe 568
Db      2452 GCGCAACTTATATTGTTTC 2472

```

```

STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72mb
COMPUTER: IBM PC compatible (NEC PC-9801 RX)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/732,242C
FILING DATE: 19910718
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPN 2-210178
FILING DATE: 10-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910532/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 972-1400
TELEFAX: (212) 370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-732-242C-8

Alignment Scores:
Pred. No.: 3 01e-216 Length: 6131
Score: 2028.50 Matches: 375
Percent Similarity: 79.72% Conservative: 77
Best Local Similarity: 66.14% Mismatches: 112
Query Match: 67.64% Indels: 3
DB: 1 Gaps: 2

US-09-904-994B-3 (1-568) x US-07-732-242C-8 (1-6131)
Qy      3 MetLysLysGlnGluLysTrpValaenThrTrpGlyProThrLysGlyAspLysValArgLeu 22
Db      989 ATGTCTCGAAAGCAATATCGGATATGTTGGACCAACTGTGCGGACCCATTCGTTG 1048
Qy      23 GlyAspThrAspLeuTrpAlaGluValGluHisAspTrpThrTrpGlyGluGluLeu 42
Db      1049 GCAGATTTCAGAAATGTTATTCGAAATTTGAAAGACTATACACGTATGAGATGAGGTA 1108
Qy      43 LysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsn---SerProAsp 61
Db      1109 AGTTTGGCGGCGGCAAGGTGATCCGAGTGAATGGGCAACATCTTGGCGACACAC 1168
Qy      62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTrpThrGlyIleTrp 81
Db      1169 GATGATGGCTGATCTGTATTAACAATGCGATATGTTGATTAACAAGATTTAT 1228
Qy      82 LysAlaAspIleGlyLysLysAsnGlyLysIleHisGlyLysGlyLysAlaGlyAsnLys 101
Db      1229 AAAGCAGATATCGCATTAAGAAAGATGATGATTCCTCATGGAAGAAAGCGGGAACCG 1288
Qy      102 AspMetGlnAspGlyValSerProHisMetValaGlyValGlyThrGluAlaLeuAla 121
Db      1289 TTGTTATGACAGCGGGTGCAT-----ATGGATTTGGAGCGACACACAGATCATAGCC 1342
Qy      122 GlyGluGlyMetIleIleThrAlaGlyLysIleAspSerHisThrHisPheLeuSerPro 141
Db      1343 GCAGAAAGGATGATGTGACACCCGAGAGATAGATGTCATATTCATCTTATTTGCCCT 1402
Qy      142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161
Db      1403 CAGCAAAATGAAACCGCTCTTCATCGGCTGTGACCACTATGATTTGGCGAGAGAACAGA 1462

```

```

QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGluLysTrpAsnLeuHisArgMet 181
Db 1463 CCCGCTACAGGCAACAAATCCACTACTGTACACGGGGCCTTGGAATATCATCGATATG 1522
QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
Db 1523 CTTCACAGCAGCCGAAGAATCCCGATTAACCTTGCGCTTTTAGGAAAGGAAACGTGTCA 1582
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221
Db 1583 GATGAGGCTCCTTTAAAGAAATGAAACGGGAGCGGCGGATTTAAACCTTCACCA 1642
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241
Db 1643 GATTGGGAGATGACGCGGGCGGCTATTGATCATGTTTGAAGTGGCGGATCATGAT 1702
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261
Db 1703 GTGCAATAGCGATTCATACAGACATTTAAATGAGCGGATTTGTCAGAGATACCTTG 1762
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGlyHis 281
Db 1763 AAAGCCATAGCGGTGATTCATACCTATACAGAAAGGCGGTGCGGCGGACAT 1822
QY 282 SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrPro 301
Db 1823 GCTCCGATATTATTAAGCGCGCGCTTCCGAAATTTTGCTTCTTCCACAGATCA 1882
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
Db 1883 ACTGACCTTATATCATCAATACCTTTGGAGAGACATTTAGATGTTATGTTTGCAC 1942
QY 322 HisLeuAspLysArgIleLeuGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 1943 CACCTACACGCTAATATTCACAGAGATATTCCTTTCGCGATTCACGATACGGAAGAG 2002
QY 342 SerIleAlaIleGluAspValLeuHisAspMetCysValIleAlaMetThrSerSerAsp 361
Db 2003 ACCATCCGCGGAGAAAGATTTTACATGATTTAGCCGTTTCCACAGATTTCCGTGAT 2062
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
Db 2063 TCCAGGCGATGGGCGGAGTAGAGAACTGATCATCTGCTACGTGCGCAACGCGTACAG 2122
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
Db 2123 ATGAAAAGCAAGAGGAAAGTTACAAAGACAAATGCTGTGGAGACAACTTCGTGTG 2182
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGluTyr 421
Db 2183 AAACGTTATATGCGAAATATACATCAATCCGCGCATTCGTCATGATGATTCGCGATTAT 2242
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTPAsnProAlaPhePhe 441
Db 2243 GTGGGTTCTGTGAAGTGGGAAATTAGCTGATTTAGTGGTGGAAATCCGCTTTT 2302
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
Db 2303 GGGTGTGAACCTGAATCGTCTTAAAGAGAGATGATGCTTACAGACATAGGAGAT 2362
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgLysMetPheGlyHisHis 481
Db 2363 CCCAATCCAGCAATTCGACACCGCACCGGTTTAAATCGTCGATGTTTGCAGCGAAA 2422
QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501
Db 2423 GGAGATCCCAAATATCAACGCTATATCACTTTTTCGAAAGACGCTATGAAAAAGGC 2482
QY 502 ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
Db 2483 ATTCATGAACAGTTGGGTGGAAGAAAGGTGAAACCAAGTCATGGAATTCGAAATTTG 2542

```

```

QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2543 ACGAAAAAGATTTAATTTTGAACGATTAACCCCAAAATTAAGTGCATCTTCAGACA 2602
QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
Db 2603 TATGAAGTAAAGGATAGACGCGTCATATGATGACATGTAACCGGACAGAAATGTCCTATG 2662
QY 562 AlaGlnArgTyrThrPhePhe 568
Db 2663 GCACACGCTAATTTCTTATTT 2683

RESULT 9
US-09-252-991A-10316
; Sequence 10316, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10316
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10316

Alignment Scores:
Pred. No.: 1,16e-194 Length: 1701
Score: 1828.00 Matches: 336
Percent Similarity: 74.47% Conservative: 87
Best Local Similarity: 59.15% Mismatches: 143
Query Match: 60.95% Indels: 2
Gaps: 1

US-09-904-994b-3 (1-568) x US-09-252-991A-10316 (1-1701)
QY 1 MetLysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
Db 1 ATGAAAATTCAGACCAAGCGTACGCCGACATGTTCCGCCCCACCGCGCGGACCGCGTG 60
QY 21 ArgLeuGlyAspThrAspLeuThrAlaGluValGluHisAspTyrThrTyrGlyGlu 40
Db 61 CGCTGGCCGACACCGGCTGTGATCGAGGTGGACGGGACCTTCAACCGTCTATGGCGAG 120
QY 41 GlnLeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro 60
Db 121 GAAGTGAAGTTCCGCGCGGCGGCAAGTCATCCGCAACGCGATGGGCCAGATCAGTGGCC 180
QY 121 AsnGlnAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80
Db 61 ACGGCGGAGGTGGCGACAGGTGATCACCAATGCGCTGATCTCGACCACTGGGGCGG 240
QY 81 TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100
Db 241 GTCAAGGCCGACGCTGCTCAAGAGCGGCGCATCCAGGCCATCCGCAAGCGCGGCAAC 300
QY 101 LysAspMetCysAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeu 120
Db 301 CCGTACATCCAGCCCGCGGCGG-----AACATCGCCATCCGCGCGCGGACGAGTATC 354
QY 121 AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140
Db 355 GCGGCGCAGGCGATGATCTTCACCGCGCGGCGATCGACACGACATCCACTTCATCTGC 414
QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr 160

```

Db	415	CCGAGGACGATCGAAGAGCGCTGATGAGCGGGGTACACCAATGATCGGGGGCAC	474
Qy	161	G1yProValaAerg1yThraAsn1aThrThr1leThrProG1yLysTrpAsnLeuH1eaArg	180
Db	475	GGCCCCGGCGCGGGAGCCAGCCACGCAACCTGACACTCCGGTCCCTGGCACATGGCGGG	534
Qy	181	MetLeuAarg1aa1ag1ug1uTyTserMetAsnValG1yPheLeuG1yLysG1yAsnSer	200
Db	535	ATGCTCCAGCGCGCGCGAGCGCTTCCCGATGAACTCGGCTTCCACCGGCAAGGGCAACCC	594
Qy	201	SerSerLysLysG1nLeuValG1nG1nValG1n1ag1yVal1leG1yPheLysLeuH1s	220
Db	595	AGCCTGCGCGCTCGCGCTGAGAGAGAGGTGGCTCGCGGCGCATCGAGCTGAAGCTGAC	654
Qy	221	G1uAspTrpG1yThrThrProSerL1a1leAspH1sCysLeuSerVal1a1AspG1uTr	240
Db	655	GAGAGACTGGGCGACACCCCGCGCGCGATCGACAATCCCTCGAGAGTGGCGAGCGGCAC	714
Qy	241	AspValG1nValCys1leH1sThraAspThraValAsnG1u1ag1yTyTValAspAspThr	260
Db	715	GACATCCAGTGGCGATCATCACCGACACCTCAACGAGTCCGCGCTTGTCGAGACACC	774
Qy	261	LeuAsn1aMetAsnG1yAArgAla1leH1sAlaTyH1s1leG1uG1yVal1ag1yG1y	280
Db	775	CTCGCGCGCTTCAAGGGCGGACCATTCACACTACACACCGAGGGGGCGCGCGCGC	834
Qy	281	H1sSerProAspVal1leThrMet1ag1yG1uLeuAsn1leLeuProSerSerThrThr	300
Db	835	CACGACCGGACATCATCAAGGCGCGCGCTTCCGCAACGTGTCTCCAGCTCGACACAC	894
Qy	301	ProThr1leProTyTThr1leAsnThraVal1aG1uH1sLeuAspMetLeuMetThrCys	320
Db	895	CCGACACCGCGCTTCAACCGCAATPCATCGACAGCACCTGACATCTGATGGTGC	954
Qy	321	H1sH1sLeuAspLysArg1leArgG1uAspLeuG1nPheSerG1nSerArg1leArgPro	340
Db	955	CACCACTTCGACCGGCGCATCGCGAGAGCGTGGCTTGGCGAGTCCGCGATCGCGCGC	1014
Qy	341	G1ySer1leA1a1ag1uAspValLeuH1sAspMetG1yVal1le1aMetThrSerSer	360
Db	1015	GAGACCATCGCGCGGAGGACATCTCCACAGACTCCGCGCGGTACAGATATCACTCC	1074
Qy	361	AspSerG1n1aMetG1yAArgAlaG1yG1uVal1leProArgThrTrpG1nThra1aAsp	380
Db	1075	GACAGCCAGCGCATGGCGGGGTGGCGAAGATACGCGCACCTTGCCACACCGCGCAC	1134
Qy	381	LysAsnLysLysG1uPheG1yLysLeuProG1uAspG1yLysAspAsnAspPheArg	400
Db	1135	AAGATGAAGGGCCAGCGCGCGCTCCACGCGCAGCGGTGGCGGACAGCAACTTCGC	1194
Qy	401	1leYsArgTyT1leSerLysTyTThr1leAsnPro1a1eLeuThrH1sG1yValSerG1u	420
Db	1195	GCCAGCGCGCTACATGCGCAAGTACCACTCAACCCGCGGATCAACCCACGCAATCCAT	1254
Qy	421	Tyr1leG1ySerValG1uG1uG1yLys1le1aAspLeuVal1a1TrpAsnPro1aPhe	440
Db	1255	GAGTGGGCTCGGTGAAGCCGGCGAAGGGGCGACCTGTGGCTTGCGGCCCTTTC	1314
Qy	441	PheG1yValLysProLys1leVal1leLysG1yG1yMetVal1aPheSerG1uMetG1y	460
Db	1315	TTCGCGCGCAAGCGAGCTATCTCTCAAGGGCGGGGCAATCGCGCAGCTGATGGCG	1374
Qy	461	AspSerAsn1aSerValProThraProG1nProValTyTArgG1uMetPheG1yH1s	480
Db	1375	GACATCAACGGCTCATCTCCACGCGCGAGCGGGTGCCTACCGCAAGATGTTGCCAGC	1434
Qy	481	H1sG1yLysAlaLysPheAspThrSer1leThraPheVal1SerLysVal1a1TyTArg1uAsn	500
Db	1435	TACGCGGCGACCGGTCAAGCGACGACGCTGACCTTGTCAGCGAAGCGCGCTTGGCGGC	1494
Qy	501	G1yValLysG1uLysLeuG1yLeuG1uArgG1nValLeuProValLysAsnCysArgAsn	520

Db	1495	GGCGTACCGCAGCAGCTCGGCGCTCGCAGAGGCCATCGCGCGTGGTGCAGCGCTGCGCGGT	155
Qy	521	ThrIyLysLysAspPheLysPheAsnAspLysThrIalValSIIeThrValAspProLys	540
Db	1555	:::GTGCAAGAAAGACCGACCTGATCCACAACGGGCTACCTACCGACCATCGAGGTGACGACAG	1614
Qy	541	ThrPheGluValPheValAspGIIyLysLeuGlySerThrSerLysProThrSerGluValPro	560
Db	1615	AACTACCAAGTACCGCGCCGACGGCGCAGTGGCTCTGTGTCGGAACCGCGGACGTGCTGCCG	1674
Qy	561	LeuAlaGlnArgTyrThrPhePhe	568
Db	1675	ATGGCGCAGCGCTATTTCCTTC	1698
	RESULT 10		
	US-08-487-429A-3		
	; Sequence 3, Application US/08487429A		
	; Patent No. 6468765		
	GENERAL INFORMATION:		
	; APPLICANT: Flischmann et al.		
	; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd		
	; Patent No. 6468765		
	; FILE OF INVENTION: Genome, Fragments Thereof, and Uses Thereof		
	; FILE REFERENCE: p186p1		
	; CURRENT APPLICATION NUMBER: US/08/487,429A		
	; CURRENT FILING DATE: 1995-06-07		
	; PRIOR APPLICATION NUMBER: 08/426,787		
	; PRIOR FILING DATE: 1995-04-21		
	; NUMBER OF SEQ ID NOS: 10		
	; SOFTWARE: PatentIn Ver. 2.1		
	; SEQ ID NO 3		
	; LENGTH: 1716		
	; TYPE: DNA		
	; ORGANISM: Haemophilus influenzae		
	US-08-487-429A-3		
	Alignment Scores:		
	Pred. No.: 1,97e-194	Length: 1716	
	Score: 1826.00	Matches: 337	
	Percent Similarity: 75.17%	Conservative: 93	
	Best Local Similarity: 58.92%	Mismatches: 136	
	Query Match: 60.89%	Indels: 6	
	Db: 3	Gaps: 3	
	US-09-904-994B-3 (1-568) x US-08-487-429A-3 (1-1716)		
Qy	1	MethylMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal	20
Db	7	TTAACAAATTTCAAGAGCCCAATATGATGACCACTTATGATCCACAAGTGGCGCATTAAGTC	66
Qy	21	ArgLeuGlyAspThrAspLeuTyrPAlaGluValGlnHisAspTyrThrThrTyrGlyGlu	40
Db	67	CGTTTAGCGCATCCAAATTTATGGGCACCAATGAAACAAGATTATATGACCAAGGTGAT	126
Qy	41	GluLeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSer---	59
Db	127	GAGGTAAATTTGGTGGCGGTAAAGCGTGGTGTATGGCTCAAAAGCGGTACGGCA	186
Qy	60	-----ProAspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIAsp---	76
Db	187	ACTGGCGCAATCCGAATGTATGGATTTGTGTGATTACCAACGGATGATCATTTGATGCT	246
Qy	77	TyrThrGlyIleTyrLysAlaAspIleGlyIleLysAsnGlyLysIleIleGlyYIleGly	96
Db	247	AAATTAGGCATTATCCAAAGCCGATATTGGTATTCGTATGGCGCGATTTGGGTATTGGA	306
Qy	97	LysAlaGlyAsnLysAspMetGlnAspGlyValSerProHisMetValIaGlyValGly	116
Db	307	CAAGCAGGTAAACCTCGACACCATGATGTACGTCACACCAAAATATGATATATGATGTCGAAC	366
Qy	117	ThrGluAlaLeuAlaGlyGluGlyMetIleIleThrAlaGlyGlyYIleAspSerHisThr	136
Db	367	ACGGAAGTTCAATTAACGGTGCACATTATTTTTCGAAACCGGTGGTGGTATGATATACCAATT	426

```

Qy 137 HispHeuSerProGInGlnPheProThraAlaLeuAlaAnglValThrThrMetPhe 156
Db 427 CACTTATTTTGTCCACACACACACACATCATTAAGAGTGGGTTACACCGTTAATT 486
Qy 157 GlyGlyGlyThrGlyProValAspGlyThrAsnAlaThrThrThrProGlyLysThr 176
Db 487 GGTCGTGGAAGTGGCCCTGCTGATGATACACCGAACCATTTGATACCCCTGGGCAATG 546
Qy 177 AsnLeuHisArgMetLeuArgAlaAlaGluGlyThrSerMetAsnValGlyPheLeuGly 196
Db 547 TATATGGAACGTATGTTCAAGCCGACAAACCTTGCCTGATTAAGCGGAGTATTTTGGT 606
Qy 197 LysGlyAsnSerSerSerLysGlyGlnLeuValGluGlnValAlaGluAlaGly 216
Db 607 AAGGCAACTGTTTACCCCTTATCCCTGCTGCTGATTAAGCGGAGTATTTTGGT 666
Qy 217 PheLysLeuHisGlyAspThrGlyThrThrProSerAlaAlaAspHisGlyLeuSerVal 236
Db 667 TTAATAATCCACGAAAGATGGGGTGCAACCGCTCCGATTTGATTTGCTTAAAGTA 726
Qy 237 AlaAspGlyThrAspValGlnValCysAlaHisThrAspThrValAsnGluAlaGlyThr 256
Db 727 GCAGATGAAATGGATATTCAGTCCCATTCACACAGACGCTAAATGAAAGTGGCTTT 786
Qy 257 ValAspAspThrLeuAsnAlaMetAsnGlyArgAlaAlaHisAlaGlyHisAlaGluGly 276
Db 787 TTGGAAAGACGAGTAAAGCATTTGATGAGACGAGCTATTCTTCCATTAACGAGAGGC 846
Qy 277 AlaGlyGlyGlyHisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuPro 296
Db 847 GCAGGTGGTGGTATGACCGTACCATTTAAAGCAGAGATGATTCAGAGTATTAACCT 906
Qy 297 SerSerThrThrProThrIleProGlyThrIleAsnThrValAlaGluHisLeuAspMet 316
Db 907 GCTTCAACCAACCCGACTCGCTCCCTTTACCAAAAACCATGATGAAACATTTGGATATG 966
Qy 317 LeuMetThrCysHisHisLeuAspLysArgIleArgIleAspLeuGlnPheSerGlnSer 336
Db 967 TTGATGGTGGTCCATCACTTATGATTAACCGCTGGGAGACGATCTTTGGCGATAGC 1026
Qy 337 ArgIleArgProGlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAla 356
Db 1027 CGTATCGCCCTGAAACATTTGACGACGAGATATTTTGCATGATATGGGCTTCTCC 1086
Qy 357 MetThrSerSerAspSerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrp 376
Db 1087 ATTATGAGTTCAGACTCTCAAGCGATGGGACGATGCGAAGTCGTTATTCGTAACATG 1146
Qy 377 GlnThrIleAspLysAsnLysGlyGlnPheGlyLysLeuProGluAspGlyLysAspAsn 396
Db 1147 CAATCTGCAATTAAGTAAATGCAACGTGTGAGCTAGGTAATGAAGA- - - - -AAC 1200
Qy 397 AspAsnPheArgIleLysArgGlyTrpIleSerLysThrIleAsnProAlaLeuThrHis 416
Db 1201 GATTAATCTCGTATTAACGATATATCGCAATATACCCATCAACCCAGCATTTGCAT 1260
Qy 417 GlyValSerGlyTrpIleGlySerValGluGluGlyLysIleAlaAspLeuValValTrp 436
Db 1261 GGTAATTCGAGACATATTTGGCTCTTAGAAGTGGTAAATCGACATATGCTGTATGG 1320
Qy 437 AsnProAlaPhePheGlyValLysProLysAlaIleValIleLysGlyGlyMetValValPhe 456
Db 1321 AAACCGATGTTCTTTGGCGTAAACCTGAAGTGTGATTAATAAAGGCTTATTAAGCAT 1380
Qy 457 SerGluMetGlyAspSerAsnAlaSerValProThrProGlnProValTrpArgGlu 476
Db 1381 GCGAAATGGCGCATTCGAATGCTCTCAATTCACACCGCACTGATTTCTACCGTCCA 1440
Qy 477 MetPheGlyHisHisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysVal 496
Db 1441 ATGTACGTCGACAAAGGCTTAGCAACGCAACAAACAGACAGATATCTTTGTTCAAGCC 1500

```

```

Qy 497 AlaTrpGluAsnGlyValLysGlyLysLeuGlyLeuGluThrArgIleValLeuProValLys 516
Db 1501 CGTAAATAAGCTGATATATTCGGCGAAGTTCGTTTACCAAGAAACCATTTGCTGAAA 1560
Qy 517 AsnCysArgAsnIleThrLysLysAspPheAsnAspLysThrAlaLysIleThr 536
Db 1561 GGTCGCGCAACGAGGTAAATAAGATCTGCTCATATGATGTAACACCAACCATTAAT 1620
Qy 537 ValAspProLysThrPheGluValPheValAspGlyLysLeuGlySerThrSerLysProThr 556
Db 1621 GTTATCTCTGAACCTTATGAAGTTCGAGTGCAGAGAGTTAATTAACCTGTGAACAGTG 1680
Qy 557 SerGlnValProLeuAlaGlnArgTrpThrPhePhe 568
Db 1681 GATACGCTACCATTTGGTGCACGATATTTCTTATTC 1716

RESULT 11
PCT-US96-05320A-541
Sequence 541, Application PC/TUS9605320A
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences
APPLICANT: 9410 Key West Avenue
APPLICANT: Rockville, MD 20850
APPLICANT: United States of America
APPLICANT: Johns Hopkins University
APPLICANT: 720 Rutland Avenue
APPLICANT: Baltimore, MD 21205
APPLICANT: United States of America
APPLICANT: Mark D. Adams
APPLICANT: Owen White
APPLICANT: Hamilton O. Smith
APPLICANT: J. Craig Venter
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20003-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05320A
FILING DATE: April 122, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Eric K. Steife
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.014PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 541:
SEQUENCE CHARACTERISTICS:
LENGTH: 1716 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US96-05320A-541

Alignment Scores: 1.97e-194 Length: 1716
Pred. No.:

```


CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.: 1,58e-189 Length: 1830121
Score: 1826.00 Matches: 337
Percent Similarity: 75.17% Conservatave: 93
Best Local Similarity: 58.92% Mismatches: 136
Query Match: 60.89% Indels: 6
Gaps: 3

US-09-904-994B-3 (1-568) X US-09-557-884-1 (1-1830121)

QY 1 MetLysMetLysbGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
Db 564173 TTAACAATTTCAGAGCGCATATGTAGCAACTTATGTCACAGCTGGCGATTAAGTC 564114
QY 21 ArgLysGluAspThrAspLysLeuValGluValGluAspLysValGluValGluVal 40
Db 564113 CGTTAGCGGATACCAATTATGAGCAACCACTTGAACCAAGTTATGACCAAAAGGTGAT 564054
QY 41 GluLeuLysPheGlyLysGlyLysThrLeuGluGluGluGluGluGluGluGluGlu 59
Db 564053 GAGGTAAATTTGGTGGCGGAAAGCGTGGTATGCTGATGCTCAAGCGCGATACGCA 563994
QY 60 -----ProAspGluAsnThrLeuAspLeuValLeuThrAsnAlaMetLeuLeuAsp--- 76
Db 563993 ACTCGACAAATCGAATGTATGATTTGTTGATTACCAAGTATGATCATTTGATGCT 563934
QY 77 TyrThrGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 96
Db 563933 AAATTAGCAATTATCAAGCGGATTTGTTGATTTGTTGATTTGTTGATTTGTTGATTTG 563874
QY 97 LysAlaGlyAsnLysAspMetGlnAspGlyValSerProHisMetValValGlyValGly 116
Db 563873 CAAGCAGTAACTCTGACACATGATTAACCTCAACCAATATGATTTATGCTGCAAC 563814
QY 117 ThrGluAlaLeuAlaGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 136
Db 563813 ACGGAACTTCAACCGGTGACATTTAATTCACCGGTGATGATGATGATGATGATGATG 563754
QY 137 HisPheLeuSerProGlnGlnPheProThrAlaLeuAlaGlyValThrThrMetPhe 156
Db 563753 CACTTATTTGTCACAAACCAACCAACATGTAAGTGGCGTTTACACCGTTAATT 563694
QY 157 GlyGlyGlyThrGlyProValAspGlyThrAsnAlaThrThrLysThrProGlyLysThr 176
Db 563693 GGTGTGGAAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 563634
QY 177 AsnLeuHisValMetLeuArgAlaAlaGluLysThrSerMetAsnValGlyPheLeuGly 196
Db 563633 TATATGAAAGTATGTTTCAAGCGGACAAAGCTTGGCGGTAAGCGTGAATTTTGGT 563574

QY 197 LysGlyAsnSerSerSerLysGluLeuValGluGluValGluAlaGlyValAlaGly 216
Db 563573 AAAGCAACTGTTACACCTTATGATCTCTGCTGATGACAAATTTGAAGCGGATGATGCT 563514
QY 217 PheLysLeuHisGlyAspTrpGlyThrThrProSerAlaLeuAspHisValLeuSerVal 236
Db 563513 TTAATAATTCACGAAGCTGGGGTGCAACGCTCCCGTATGATTTATGCTTAAAGTA 563454
QY 237 AlaAspGluTyrAspValGluValCysLeuHisThrAspThrValAsnGluAlaGlyTyr 256
Db 563453 GCATATGAATATGATTTCAAGTGGCCATTCACACACACGCAATTAAGAAAGCGCTTT 563394
QY 257 ValAspAspThrLeuAsnAlaMetAsnGlyArgAlaLeuHisAlaTyrHisAlaGlyGly 276
Db 563393 TTGAAAGACACAGTAAAGCAAGCATTTGATGACAGATCTTCTATCTTCCATTCAGAGGCT 563334
QY 277 AlaGlyGlyGlyHisSerProAspValLeuThrMetAlaGlyLysLeuAsnLeuPro 296
Db 563333 GCAGTGTGTGATGACACGTCGATCATTTAAGCAGATGATTCACACGATTTAAGCT 563274
QY 297 SerSerThrThrProThrLysProTyrThrLysAsnThrValAlaGluHisLeuAspMet 316
Db 563273 GCTTCAACCAACCGACTCGCTCCGTTTACCAAAACCAATTTGATGAAACATTTGGAATG 563214
QY 317 LeuMetThrCysHisHisLeuAspLysArgGluLeuGluAspLeuGlnPheSerGlnSer 336
Db 563213 TTGATGTTTGGCAATCTATGATTAACAGGATGCGGAGACGATGCTTTGGCGATAGC 563154
QY 337 ArgLeuArgProGlySerLeuAlaAlaGluAspValLeuHisAspMetGlyValAlaAla 356
Db 563153 CGTATCCGCTGAAACCATTTGACGACAGAAAGATATTTGATGATATGAGGCTTCTCC 563094
QY 357 MetThrSerSerAspSerGlnAlaMetGlyArgAlaGlyGluValLeuProArgThrTrp 376
Db 563093 ATTATGATTTGACAGCTCTCAAGCATGAGAGCTATTTGCGCAAGCTTATTCGATCAATG 563034
QY 377 GluThrAlaAspLysAsnLysGluPheGlyLysLeuProGluAspGlyLysAspAsn 396
Db 563033 CAACCTCAGATTAAGTAAAGTAAAGTCAACGTTGATGCTAGGTAATGAAGG-----AAC 562980
QY 397 AspAsnPheArgLysLysArgTyrLysSerLysTyrThrLysAsnProAlaLeuThrHis 416
Db 562979 GATTAATCTTCGATTAACGATATATCGCAAAATACACATCAACCGACGATTCGACAT 562920
QY 417 GlyValSerGluTyrLysGlySerValGluGluGlyLysLeuAlaAspLeuValValTrp 436
Db 562919 GGTATTCGGAAGCATTTGGCTCGTTAGAAAGTGGTAAATTCGACATATCGTGTATGG 562860
QY 437 AsnProAlaPhePheGlyValLysProLysAlaValLeuLysGlyLysMetValValPhe 456
Db 562859 AAACCGATGTTCTTTGGCGTAAACCTGAAGTGTGATTAATAAAGCTTTATTAAGCTAT 562800
QY 457 SerGluMetGlyAspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGlu 476
Db 562799 GCGAAATGGCGCATCCAAATGCTCAATTCCAACACCGACCTGATTTCTACCGTCA 562740
QY 477 MetPheGlyHisHisGlyLysValAlaLysPheAspThrSerLeuThrPheValSerVal 496
Db 562739 ATGACAGCTGACAAAGCTTACGAAACCGCAAAACAGACAGTATCTTTGTTTACAGACC 562680
QY 497 AlaTyrGluAsnGlyValLysGluLysLeuGlyLysLeuGluGlnValLeuProValLys 516
Db 562679 GCTGAATAACCTGATATTCGTGCGAATTCGTTTAAACAAAGAACCATTTGCTGTAATA 562620
QY 517 AsnCysArgAsnLeuThrLysLysAspPheLysAspLysThrAlaLysLysLeuThr 536
Db 562619 GCGTCCGCGCAACGATGATTAATAAAGATCTGTTTATATATGATTAACCAACATTAAT 562560
QY 537 ValAspProLysThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThr 556
Db 562559 GTTATGCTGTAACCTTATGAAGTTTCGAGTGAACGAGAGATTTAATTAATCTGAACAGT 562500
QY 557 SerGlnValProLeuAlaGlnArgTyrThrPhePhe 568

```

Db      562499  GATAGCGTACATTGGGTGACGATATTTCCTATTTC 562464
                                         |||
RESULT 13
US-09-643-990A-1/c
: Sequence 1, Application US/09643990A
: Patent No. 6528289
: GENERAL INFORMATION:
: APPLICANT: Robert D. Fleischmann
:           Mark D. Adams
:           Owen White
:           Hamilton O. Smith
:           J. Craig Venter
: TITLE OF INVENTION: The Nucleotide sequence of
:                   the Haemophilus influenzae Rd Genome, Fragments
:                   Thereof, and Uses Thereof
: NUMBER OF SEQUENCES: 1
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville,
: STATE: MD
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3 1/2 inch diskette
: COMPUTER: Dell Pentium
: OPERATING SYSTEM: MS DOS V6.22
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/643,990A
: FILING DATE: 23-Aug-2000
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/487,429
: FILING DATE: 1995-06-07
: APPLICATION NUMBER: 08/426,787
: FILING DATE: 1995-04-21
: ATTORNEY/AGENT INFORMATION:
: NAME: Kenley K. Hoover
: REGISTRATION NUMBER: 40,332
: REFERENCE/DOCKET NUMBER: PB186P1C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 301-610-5790
: TELEFAX: 310-309-8439
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1830121 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
:
:
US-09-643-990A-1
:
:
Alignment Scores:
Pred. No.:
1.58e-189      Length: 1830121
1826.00      Matches: 337
Percent Similarity: 75.17%      Conservative: 93
Best Local Similarity: 58.92%      Mismatches: 136
Query Match: 60.85%      Indels: 6
DB: 4      Gaps: 3

US-09-304-994B-3 (1-568) x US-09-643-990A-1 (1-1830121)
OY      1      MetyswMetysblysglmglyurrtalantthrrtyrglyprrothrtlyeglyasplysval 20
      :      :      :      :      :      :      :      :      :      :      :
Db      564173  TTAACAAATTCAAGAGCGCCAAATATGTATGACACTTATAGTCCAAACAGTTGCGCATTAAGATC 564114
      :      :      :      :      :      :      :      :      :      :      :
OY      21      ArgLeuGIYAaPRTThAspLeuTrPAaIguValaGluNH:AspTyTrThrTYrGIyGlu 40
      :      :      :      :      :      :      :      :      :      :      :
Db      564113  CGTTTAGCGCATCAATTTATTATGGCGCAACCATTTGAACAAGATTTATTGACCAAAAGTGAT 564054
      :      :      :      :      :      :      :      :      :      :      :
OY      41      GluLeuYsphegiYlaGIyLysThrIleatrgGluGIyMeGIyGlnSerAnSer--- 59
      :      :      :      :      :      :      :      :      :      :      :

```

Db 564053 GAGTGTAAATTTGGTGGCGGTTAAAGCGTGGCTGATGTGGCTCAAGCGGTACGGCA 563994

Qy 60 -----ProAspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAsp--- 76

Db 563993 ACTCGCGAACAATCCGAATGTATTTGATTTTGATTTACCAAGCGTATGATCATTTGATGCT 563934

Qy 77 TyrThrGlyIleTyrLysAlaAspIleGlyIleLysAsnGlyLysIleGlyIleGly 96

Db 563933 AAATTGGCATTTATCAAGCCCATTTGGTATTTCCGTAATGGGCGTATTTGGGTATTTGGA 563874

Qy 97 LysAlaGlyAsnLysAspMetClnAspGlyValIserProHisMetValGlyValGly 116

Db 563873 CAAAGAGGTAAACCTTGACCCATGAGATTAACGCACCAAAATATGATTTATCGGTGCAAGC 563814

Qy 117 ThrGluAlaLeuAlaGlyLysGlyMetIleIleThrAlaGlyIleLysAspSerHisThr 136

Db 563813 ACGGAAGTTCATACAGGTGCACATTTAATTTGCAACCGCTGGTGGTATTCGATACCCACATT 563754

Qy 137 HisPheLysSerProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPhe 156

Db 563753 CACTTATTTTGTCCACAACAGCAACAATGCAATTAAGAGGCGCTTACACAGTTAAAT 563694

Qy 157 GlyIleGlyIleThrGlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrp 176

Db 563693 GGTGGTGAACCTGGCCCTGGCTGATGTACACAGCAACCATTTGATCCCTGGCGCATGG 563634

Qy 177 AsnLeuHisArgMetLeuArgAlaIleGluGlyIleLysSerMetAsnValGlyPheLeuGly 196

Db 563633 TATATGGAACGATATTTTCAACCGGCAAGACCTTGGCGGTAAACGTCCGATTTTGGT 563574

Qy 197 LysGlyAsnSerSerSerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGly 216

Db 563573 AAAGGCAACTGTTCACACCTGATCTCTGCTGAGCAAAATGAAGCGGTGATTAAGT 563514

Qy 217 PheLysLeuHisIleGluAspTrpGlyIleThrThrProSerAlaIleAspHisCysLeuSerVal 236

Db 563513 TTAAATAATCCAGAAAGCTGGGGTGCACAGCTGCGGTGATTTGATTTCTCTTAAAGTA 563454

Qy 237 AlaAspGlyIleLysAspValGlnValCysIleHisThrAspThrValAsnGlnAlaGlyLys 256

Db 563453 GCAAGTGAATATGATATTTCAATGGCCATTTCACAGACACCGCTAATTAATGAATGGCTTT 563394

Qy 257 ValAspAspThrLeuAsnAlaMetAsnGlyArgAlaIleHisAlaIleTyrHisIleGlyIle 276

Db 563393 TTGAAGAACAAGATGAAGAAGCATGATGAGACGAGTCATTATTAATCTTCATACCGGAGGCG 563334

Qy 277 AlaGlyIleGlyHisSerProAspValIleThrMetAlaGlyIleLeuAsnIleLeuPro 296

Db 563333 GCAGGTGGTGGTCATGCACCTGACATCACTTAAAGCAGCATGATTTCAACAGTATTACCT 563274

Qy 297 SerSerThrThrProThrIleProTyrThrIleAsnThrValAlaGluHisIleAspMet 316

Db 563273 GCTTAAACCAACCCGACCTGCTCCGTTTACCAAAACACCATTTGATGAACATTTGGATATAG 563214

Qy 317 LeuMetThrCysHisHisIleuAspLysArgIleArgGlyAspLeuGlnPheSerGlnSer 336

Db 563213 TTGATGGTTTGGCATCACTTGATTAAGCGCGCGGAAGACGTAGCTTTGGCCATACG 563154

Qy 337 ArgIleArgProGlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAla 356

Db 563153 CGTATCCCGCTGAAACCATTCGACGACAAAGATATTTTGGCAAGATAGGCGCTTCTTCC 563094

Qy 357 MetThrSerSerAspSerGlnAlaMetGlyArgAlaGlyValIleProArgThrTrp 376

Db 563093 ATTATGATGTCACTCTCAACGCGAGGACGATTTGGGAAAGTGGTATTTGATCATGG 563034

Qy 377 GlnThrAlaAspLysAsnLysLysGlnPheGlyLysLeuProGlnAspGlyLysAspAsn 396

Db 563033 CAATGTCAAGTATAGATGAATAATGACAGTGGTGAAGCTATCGTAATGAAGCA-----AAC 562980

Qy 397 AspAsnPheArgIleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHis 416

```

Db 562979 GATTAATCCGATTAAACGATATATCGCAAAATACACCATTACACCAGCAATTCACAT 562920
Qy 417 GYValSerGIuTYrIIeGISeValGIuGIuGIyLSIleAlAspLeuValITp 436
Db 562919 GGTATGGCGAGCATATTCGCTCGTTAGAACGGGTTAAATGCGATATCGTTATAG 562860
Qy 437 AsnProAlaPhePheGIyValIlePheProLySileValIIleYsgIyGIyMeValIle 456
Db 562859 AAACCGATGTTCTTGGCGTAAACCTGAGTGGATTAATAAAAGGCTTATTGCTAT 562800
Qy 457 SerGIuMeGIyAspSerAsnAlaSerValProThrProGlnProValITyTYrArgIu 476
Db 562799 GCGAAATGGCGAGTCAAAATGCTCAATTCACACCGCAACCTGATTTACCGTCCA 562740
Qy 477 MetPheGIyHISIGLYLeAlaIleAspPheAspThrSerIleThrPheValSerIleVal 496
Db 562739 ATGTACGCTGCACAAAGCTTAGCAACCCGCAACACAGCATTTCTTTGTTTACAAAGCC 562680
Qy 497 AlaTYrGIuAsnGIyValIleGIyLSIleuGIyLeuGIuArgIuValIleuProValIyS 516
Db 562679 GCTGAAAAGCTGATATTCTGTCGCAAGTTCCGTTTACCAAAAGAACATTGCTGAAA 562620
Qy 517 AsnCySAlaGlnIleThrIleYleYleAspPheAspPheAsnAspLySThrAlaLySIIeThr 536
Db 562619 GGCTGCCGCAACGTAAAGTAAAGATCTGTTCAATGATGTAAACCAACCAATTAAT 562560
Qy 537 ValAspProLySThrPheGIuValIlePheValAspGIyLSIleuCySThrSerIleProThr 556
Db 562559 GTTATGCTGAACGTATGAAAGTTGAGTGGACGGAGGTTAATTAATCTGTGAACCAAGTG 562500
Qy 557 SerGIuValProLeuAlaGlnIleAspTYrThrPhePhe 568
Db 562499 GATAGCGATACCATTTGGGTCACGCAATATTCCTATTC 562464

RESULT 14
US-09-489-039A-2045
; Sequence 2045, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2045
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2045

Alignment Scores:
Pred. No.: 8.3e-194 Length: 1878
Score: 1821.00 Matches: 338
Percent Similarity: 74.36% Conservative: 84
Best Local Similarity: 60.04% Mismatches: 139
Query Match: 60.72% Indels: 2
Gaps: 1

US-09-904-994b-3 (1-568) x US-09-489-039A-2045 (1-1878)
Qy 3 MetIleValSerGIuTYrIIeGISeValGIuGIuGIyLSIleAlAspLeuValITp 22
Db 13 ATTTCAGCGCAGGCTATGCGCATATGTTCCGCCCAACCGCGGCAAGATGCGGCTG 72
Qy 23 GYAspPheAspLeuValITpAlaGIuValIlePheAspTYrThrIleYleYleGIuLeu 42
Db 73 GCGGATACCGGAGCTGTGATCGAGTGAAGCATATCGACCACTCGGGGAAAGGCTC 132
Qy 43 LySAspPheGIuValIleGIyLSIleArgGIuGIuGIyMeGIyIleSerAsnSerProAspIu 62

```

```

Db 133 AAATTCGCGCGCGCAAGGTATCGCGACCGCATGGGCGACGGGCGACAGATCGGCGCC 192
Qy 63 AsnThrIleAspLeuValIleThrAsnAlaMetIleIleAspTYrThrGIyIleYrIyS 82
Db 193 GACTGTGTGCACCTGTGTCTCACCAACGCTTGATGTGATGTCATCGGGGATCGTTAAG 252
Qy 83 AlaAspIleGIyIleYleYleAsnGIyLSIleGIyLSIleGIyLSIleAlaGIyAsnIleAsp 102
Db 253 GCCGATTCGCGCGTGAAGACCGCGCGATCTTCGCCCATCGCGCAAGCGGTTAACCCGAT 312
Qy 103 MetGIuAspGIyAlaSerProHISMeValValGIyAlGIyThrGIuAlaLeuAlaGIy 122
Db 313 ATCCAGGCCCAACGTACAC-----ATCCCATCGGTGCTCGACGGAAGTATGCGCGCC 366
Qy 123 GIuGIyMeTIleIleThrAlaGIyGIyIleAspSerHISThrHISPheIleSerProGln 142
Db 367 GAAAGAAAATGTGCACCGCGCGGAGTGCATCCCATATTCATCGATGTGTCCGAG 426
Qy 143 GlnPheProThrAlaLeuAlaAsnGIyValIleThrIleMetPheGIyGIyIleThrGIyPro 162
Db 427 CAGCGGAAAGCGGCTGTGCTCGGCGTACACCATGTGTGCGCGGCGGCGGCGCGCGCG 486
Qy 163 ValAspGIyThrAsnAlaThrThrIleThrProGIyIleSTraAsnLeuHISArgMetIleu 182
Db 487 GCGCGGCGCACCATGCGCACACCATGCAACCCGCGCGCGTGTATCTCATCGCATGCTG 546
Qy 183 ArgAlaAlaGIuGIyTYrSerMetAsnValGIyPheIleuGIyLSIleYleAsnSerSer 202
Db 547 CAGCGCGCGCACCGCTCGCATATTCGCTGTGCGGCGGCAAGGCGCAAGTTTCTGAG 606
Qy 203 LySAspGIuLeuValGIuGIuValIleGIyAlaIleGIyPheYleuHISGIyAsp 222
Db 607 CCGAGTGCCTCGCGCGACAGGTGCGGCGGCGCTTATGCGCTGAAGATCATGAGGAC 666
Qy 223 TrpGIyThrThrProSerAlaIleAspHISCySLeuSerValAlaAspGIyTYrAspVal 242
Db 667 TGGGGCGGACCGCGGGGCGATGCACTGTCGTTAACCGTCCGCGCATGAATGAGATC 726
Qy 243 GlnValCySIIeHISThrAspThrValAsnGIuAlaGIyTYrValAspAspThrIleuAsn 262
Db 727 CAGGTGCGCGTGCACACGACACCATTAATGATCCGCTTTGTGGAAACACACCTCGCC 786
Qy 263 AlaMetAsnGIyAspAlaIleHISAlaTYrHISIleGIuGIyAlaGIyIleYleSer 282
Db 787 GCGATCGCGGCGCGCACCATTCACACTTCATACCAAGGCGCGCGGCGGCGCATGCG 846
Qy 283 ProAspValIleThrMetAlaGIyIleuAsnIleuProSerSerThrThrProThr 302
Db 847 CCGGACATCATCACCGCTGCGCGCCACCCGCAACATTTTGCCTGTGTCCACCAACCCAAAG 906
Qy 303 IleProTYrThrIleAsnThrValAlaGIuHISleuAspMetIleuMetThrCySIIeHIS 322
Db 907 CTGCGCTACACCTGCACACCATCATGATGATCATGATGCTGATGCTGTGCGACCAT 966
Qy 323 LeuAspIleArgIleArgGIuAspLeuGIuPheSerIleAsnArgIleArgProGIySer 342
Db 967 CTGACCGCGACATCGCGGAGAGGTGCTTGCAGAGTGCAGATTCGCGGGAAC 1026
Qy 343 IleAlaIleAspValIleHISAspMetGIyAlaIleAlaMetThrSerSerAspSer 362
Db 1027 ATGCGTGGAAAGACGCTGCTGCACGATCTGCGTCTGTCTGCTGCTGCTGCTGCTGCTG 1086
Qy 363 GlnAlaMetGIyArgAlaGIyIleValIleProArgThrTYrGIuIleThrAlaAspIleAsn 382
Db 1087 CAGGCGATGGGCGCGTGGGGAAGTATCTCGGACCTGCGACGTGGCGGCGCATCGCATG 1146
Qy 383 LySAspIlePheGIyLSIleuProGIuAspGIyIleAspAsnAspAsnAspAsnAspAsn 402
Db 1147 AAGGTGACGCGGAGCGTGGCGGAGAGACCGGGATTAACCAACTTCGCGTGAAG 1206
Qy 403 ArgTYrIleSerIleTYrThrIleAsnProAlaIleuThrHISGIyValSerGIyIle 422

```

Db	1207	CGTACATCGCAAAATACACCATACACC	CGGCGCTGACCCAGCATGACAGAACT	1266	
Qy	423	GLYSERVALGIUGIUGLYLSEI	LAASPLEUVALITRPAENPROALPHEPHEGLY	442	
Db	1267	GGATTCATTGGGTGAGGTAAAGCTGGCTGA	CCCTCGGTGTGTGTACACGACCTTCTTGCGC	1326	
Qy	443	VALYSPROLYSILEVALILEYSGI	GLYMETVALVALPHESENGIUMETGLYASPSER	462	
Db	1327	GTEAAACCGCGCACCGTGATCAAAAGCGCGCAT	GTGCGCATCGGCGCGCATGGCGCATTC	1386	
Qy	463	ASNAIASERVALPROTHIRPROGILNPROVALIT	YTRYARGLUMETPHEGLYHISGLY	482	
Db	1387	AATCCCTTATTCGCAACCCCGACGCGGTGAC	TACCGCCGAGTGTGGCGCGCTGGC	1446	
Qy	483	LYSALYALYSPHEAPTHSERILETHR	PHAYLESLYVALALATYRGILNAGIYVAL	502	
Db	1447	AGCGCCGCGCATCTACCTGCCCTTCACTCT	CTGTGTGACAGCGCGGAGACCCAAATGGCCTT	1506	
Qy	503	LYSGILNYSLEUGIYLEUGILNARGILNAL	LEUPROVALIYSAENCYARASNTILETHR	522	
Db	1507	GCCGACGCGGTGAACCTGCGCAGCGCATTC	CCCGTGGAAAGGTGTGCTACGCTGACG	1566	
Qy	523	LYSLYSAAPRHEIYSPHEASNAAPLYS	THRALALYSILETHRVALASP	PROLYSTHRPHE	542
Db	1567	AAACCCGACATGTGACACACAGCCTTGAC	GCCTTAACATCACCGTGCAGCCCAAGCCTAT	1626	
Qy	543	GLIUVALPHEVALASPGIYLYSLEUCY	THRSERIYSPROTHIRSERGINVALPROLEU	VAL	562
Db	1627	GAGGTGCGGGTGGATGTGGCAACTTAT	CAACGACGCGGAGACAGTCGTCGCGATGGCG	1686	
Qy	563	GLINARGLYR	565		
Db	1687	CAACGATAT	1695		

Db 124 GCGCAAGAGTCAAATTTGTGTGGTGAAGTATTTCTGATGGATGGGCAACAGCCAA 183
|||:::|||||
Qy 59 SerProaspGlnAsnThrLeuaspLeuValIleThrAsnAlaMetIleLeaspTyrThr 78
:::|||||
Db 184 GTTGTAGTCGTCAGTGTGTCGATGTTGTCGACCAATGGCATTATTTTGAATTATTTGG 243
|||:::|||||
Qy 79 GlyIleTyrIleValAlaSprIleGlyIleTyrAsnGlyLysIleHisGlyIleGlyLysAla 98
|||:::|||||
Db 244 GGCATTGTAAACCAATATTTGGCATTTAAGAATGGCCGATATTTGCGGATTTGGCAAGCA 303
|||:::|||||
Qy 99 GlyAsnLysAspMetGlnAspGlyValSerProHisMetValGlyValGlyThrGlu 118
:::|||||
Db 304 GGTAATTCAGATTTTCAGCCCAATGGAT-----ATTGCATTTGGCCCGGAAACGAA 357
|||:::|||||
Qy 119 AlaLeuAlaGlyGlnGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPhe 138
|||:::|||||
Db 358 GTTGTTCGGGAAAGATAAATTTGTCATGCTGGTGGTATTTGATACCCATATCCACTTT 417
|||:::|||||
Qy 139 LeuSerProGlnIlePheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGly 156
:::|||||
Db 418 ATTTGTCGGCAACAAACCCCAAGAGGTCGTTTCTGGCGATAACACCTTATTTGGTGA 477
|||:::|||||
Qy 159 GlyThrGlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTyrAsnLeu 178
|||:::|||||
Db 478 GGAACAGGCCCTTGGCGGGTACTAATGCAACACAGTTACCCCGGATTTGGAAATATG 537
|||:::|||||
Qy 179 HisArgMetLeuArgAlaAlaGlnGlyIleTyrSerMetAsnValGlyPheLeuGlyLysGly 198
|||:::|||||
Db 538 TAACGCATGTTAAGCGCGGTAGTAAATTAAGTATTAATGAGGGGTATTTGTAAAGGT 597
|||:::|||||
Qy 199 AsnSerSerSerLysGlnLeuValGlnGlnValGlnAlaGlyAlaIleGlyPheLys 218
|||:::|||||

```

RESULT 15
US-09-543-681A-1857
Sequence 1857, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 1857
LENGTH: 1710
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-1857

```

QY	219	LeuHISgluAspHrTgcGlyThrThrProserAlaIleAspHIScYSLeuSerValAlaAsp	233

Db	658	ATGATATTAAGACTGGGGGGGCAACGCCAAATGGCAATTCACAATTGCTTAAATGTGGCCGAT	711
QY	239	GIuTYrAspValGIuValcYSIleHISThrAspThrValaAsnGIuValaGIYTYrValaAsp	255

Db	718	GAATGATATGTCAGGGTGCTATTCTGTGCACCTTAATTAAGAGGGTGTTTTATGA	777
QY	259	AspThrIleuAsnAlaMetcAsnGIyATgAlaIleHISAlaTYrHISIleGIuGIYValaGIY	275

Db	778	GAGCACTATAAACCCCATTCGCGGTGCAGTATCCATGTATTCCAAACCGAAGGGCGAGGT	837
QY	279	GIYGIYHISerProAspValIleThrMetAlaGIYGIuLeuAsnIleLeuProSerSer	295

Db	838	GGTGGTATGATCCCCCTGATGTATCAAGTCGGTAGAGAGGCCCAATTTTACCTGCATCA	897
QY	299	ThrThrProThrIleProTYrThrIleAsnThrValAlaGIuHISLeuAspMetLeuMet	315

Alignment Scores:	
Pred. No.:	1,366-193
Score:	1818.50
Percent Similarity:	75.44%
Best Local Similarity:	59.65%
Query Match:	60.64%
DB:	4
	2

QY 319 ThcYshiststrenuaplsarsgllleArglnuapleanglnpneseglnserArglle 336
 |||||
 Db 958 GTCGTGATCATCCGATCCCTCTATTCCTGAAGATGGCATTTCGCAATTCGATT 101
 |||||
 QY 339 ArgProciyserllelaaaglnupValleuhsaspmetgVallelaameThr 356
 |||||

US-09-904-994B-3 (1-568) x US-09-543-681A-1857 (1-1710)

QY 2 LysMetLys-----LysGInGluIuYValAsnThrIYrGlyProThrLysGlyAsp 18

Db 4 AAAATGAAACTATCTTCACGTCACGCTTATGCGGATATGTTGGCCCAACACAGGCAT 63

QY 19 LysValArgLeuGluYAspThrAspLeuIleuTrrAlaGluValGluHisAspTyrThrTyr 38

Db 64 CGTTTGGCATTTAGCAGATACCGACGCTGTTCTTGAATGAAAAAGATTTTACACCTTAT 122

QY 39 GlyLcIuGluIuLeuYsPheGlyAlaGlyLysThrIleArgGluGlyMetGlyLcIuIserAsn 58

QY 359 SerSerAspSerGlnAlaMetGlyArgAlaGlyIleVal111LeProArgThrTyrGlnThr 378
 Db 1078 TCCTGCAGACTCACAGGCATGGGACGAGCGAGAAATTACTTACGCACTTGCACAGTGT 1123
 QY 379 AlaAspLysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsn 398
 Db 1138 GCACATAAATGAATTGCACAGCGCATTTAGCGGGTGATAGCGCAGATAATGATATAT 1143
 QY 399 PheArgLleLysArgTyrLleSerLysTyrThrIleAsnProAlaLeuThrHisGlyVal 411

```
Db 1198 AATCGATTAAACGTTATATGCTAAATACAGATTAAATCCCGCACTGGCAATGGCAATT 1257
QY SerGluTyrIleGlySerValGluGluGlyLeuIleAlaAspLeuValIleTyrAspPro 438
Db 1258 GCTCATACGGTGGAGTCAATGAAAGGTAACCTTGGATATCGATGCTATAGGATCCT 1317
QY AlaPhePheGlyValIlePheProIleValIleLeuGlyGlyMetValIlePheSerGlu 458
Db 1318 GCTTCTTTGGCGTCAAAACCGCACTTATCATAAAGGTGATGCTATATCGGCCA 1377
QY MetGlyAspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPhe 478
Db 1378 ATGGGGGATATTAAATGCGCTATTCCAACCGCAACCGGTTCAATTATCGTCCAATGTAT 1437
QY GlyHisIleGlyLeuAlaIlePheAspPheSerIleThrPheValSerIleValAlaTyr 498
Db 1438 GCTGTTTAAAGGAAAGCCAAATATCAACGTCGATGATCTTATGTCAAAAGCGGTATT 1497
QY GluAsnGlyValIleGlyLeuGlyLeuGluArgGlnValLeuProValIleAsnCys 518
Db 1498 GAGCGGGAGTCCAGAAATTAAGCTTAATAAGCTTAATTGTCGTGAGGCGCTGT 1557
QY ArgAsnIleThrLeuIleAspPheIlePheAsnAspIleThrAlaIleIleThrValAsp 538
Db 1558 CGTCATATCACAAAAGCTTCGATGATCCACATATCAATGTTCTCATATCGAAATTAGAT 1617
QY ProIleThrPheGluValIlePheValAspGlyLeuGlyLeuCysThrSerIleProThrSerGln 558
Db 1618 CCACAAACTTACATGTTAAAGCGATGTTATACCACTGTTGTGAGCGACGACGTAA 1677
QY ValProLeuAlaGlnArgTyrThrPhePhe 568
Db 1678 TTACCGATGGCTCAACGCTATTCTTATT 1707
```

Search completed: February 16, 2005, 01:31:21
Job time : 2119.33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2005, 20:34:30 : Search time 8.53904 Seconds
(Without alignments)
1975.712 Million cell updates/sec

Title: US-09-904-994B-2

Sequence: 1 VKLTPKEQEFLLIYAGEVA.....ADBRHKALDKAKSHQFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 324380

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTUS COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281.5	23.8	100	1 US-08-967-513-3	Sequence 3, Appl1
2	281.5	23.8	100	2 US-08-687-645B-3	Sequence 3, Appl1
3	269.5	22.8	99	1 US-07-732-342C-1	Sequence 1, Appl1
4	261.5	22.1	100	2 US-08-467-822-23	Sequence 23, Appl1
5	261.5	22.1	100	3 US-08-432-697-23	Sequence 23, Appl1
6	261.5	22.1	100	3 US-08-466-248-23	Sequence 23, Appl1
7	250.5	21.2	100	4 US-09-602-777A-18	Sequence 18, Appl1
8	68.5	5.8	40	4 US-09-602-777A-12	Sequence 12, Appl1
9	62	5.2	97	4 US-09-640-211A-973	Sequence 973, App
10	60	5.1	63	4 US-09-248-796A-27427	Sequence 27427, A
11	58	4.9	18	4 US-08-945-038-12	Sequence 12, Appl1
12	57	4.8	92	4 US-09-270-767-60797	Sequence 60797, A
13	57	4.8	93	4 US-09-513-999C-4321	Sequence 4321, A
14	57	4.8	100	4 US-09-328-352-7014	Sequence 7014, Ap
15	56.5	4.8	83	4 US-09-248-796A-23863	Sequence 23863, A
16	56	4.7	15	3 US-09-081-001-1	Sequence 1, Appl1
17	56	4.7	64	4 US-09-328-352-6741	Sequence 6741, Ap
18	56	4.7	86	4 US-09-621-976-4833	Sequence 4833, Ap
19	55.5	4.6	100	4 US-09-621-976-7112	Sequence 7112, Ap
20	54.5	4.6	94	4 US-09-270-767-36087	Sequence 36087, A
21	54.5	4.6	94	4 US-09-270-767-51304	Sequence 51304, A
22	54	4.6	72	4 US-09-540-236-3389	Sequence 3389, Ap
23	54	4.6	85	4 US-09-248-796A-28048	Sequence 28048, A
24	53	4.5	82	4 US-09-689-065B-8	Sequence 8, Appl1
25	52	4.4	77	3 US-08-936-165A-511	Sequence 511, App
26	51.5	4.4	65	4 US-09-248-796A-26949	Sequence 26949, A
27	51.5	4.4	68	4 US-09-107-433-4410	Sequence 4410, Ap

28	51.5	4.4	72	4 US-09-512-563C-61	Sequence 61, Appl1
29	51.5	4.4	81	1 US-08-225-757B-8	Sequence 8, Appl1
30	51	4.3	66	4 US-09-732-210-881	Sequence 881, App
31	51	4.3	100	4 US-09-540-236-3479	Sequence 3479, App
32	50.5	4.3	80	4 US-09-252-991A-25293	Sequence 25293, A
33	50.5	4.3	90	4 US-09-902-540-10059	Sequence 10059, A
34	50.5	4.3	93	4 US-09-248-796A-26429	Sequence 26429, A
35	50.5	4.3	94	3 US-09-147-550-81	Sequence 81, Appl1
36	50.5	4.3	94	3 US-09-557-817-81	Sequence 81, Appl1
37	50	4.2	86	4 US-09-621-976-6687	Sequence 6687, App
38	49.5	4.2	45	4 US-09-079-030-112	Sequence 112, App
39	49.5	4.2	85	4 US-09-134-000C-3459	Sequence 3459, App
40	49.5	4.2	86	4 US-09-543-681A-5320	Sequence 5320, App
41	49.5	4.2	94	3 US-09-147-550-38	Sequence 38, Appl1
42	49.5	4.2	94	3 US-09-557-817-38	Sequence 38, Appl1
43	49.5	4.2	95	4 US-09-513-999C-5808	Sequence 5808, App
44	49	4.1	79	4 US-09-513-999C-6886	Sequence 6886, App
45	49	4.1	83	3 US-09-138-721-19	Sequence 19, Appl1

ALIGNMENTS

RESULT 1
US-08-967-513-3
Sequence 3, Application US/08967513
Patent No. 5783436
GENERAL INFORMATION:
APPLICANT: Robert P. Hausinger
TITLE OF INVENTION: Mutant Urease and Method
TITLE OF INVENTION: of Use For Determination
TITLE OF INVENTION: of Urea
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,513
FILING DATE: 11-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/687,645
FILING DATE: July 26, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 57834366
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 100
TYPE: amino acids
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Protein
DESCRIPTION: NO
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: N/A

STRAIN: N/A
INDIVIDUAL ISOLATE: N/A
CELL TYPE: N/A
FEATURE:
NAME/KEY: subunit Urea
LOCATION:
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION: encoded subunit of
OTHER INFORMATION: mutant urease
US-08-967-513-3

Query Match 23.8%; Score 281.5; DB 1; Length 100;
Best Local Similarity 59.0%; Pred. No. 4e-24;
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGVARRKKAEGKLNQPEALAIYISAHIMDEARCKTTVAQLMEE 60
DB 1 MELTPREKDKLLFTALVAERRLRGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59
QY 61 CMHPLKDEVMPCVGMV/PDLGVEATFPDGTGLVTNNMPI 100
DB 60 GRHVLTRQVMEGVPEMIPDIOVEATFPDGSKLVTNNMPI 99

RESULT 2
US-08-687-645B-3
Sequence 3, Application US/08687645B
Patent No. 5846752

GENERAL INFORMATION:
APPLICANT: Robert P. Hausinger
TITLE OF INVENTION: Mutant Urease and Method
TITLE OF INVENTION: of Use For Determination
TITLE OF INVENTION: of Urea
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,645B
FILING DATE: July 26, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5846752e
INFORMATION FOR SEQ. ID NO. 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 100
TYPE: amino acids
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Protein
DESCRIPTION: No
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:

ORGANISM: N/A
STRAIN: N/A
INDIVIDUAL ISOLATE: N/A
CELL TYPE: N/A
FEATURE:
NAME/KEY: subunit Urea
LOCATION:
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION: encoded subunit of
OTHER INFORMATION: mutant urease
US-08-687-645B-3

Query Match 23.8%; Score 281.5; DB 2; Length 100;
Best Local Similarity 59.0%; Pred. No. 4e-24;
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGVARRKKAEGKLNQPEALAIYISAHIMDEARCKTTVAQLMEE 60
DB 1 MELTPREKDKLLFTALVAERRLRGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59
QY 61 CMHPLKDEVMPCVGMV/PDLGVEATFPDGTGLVTNNMPI 100
DB 60 GRHVLTRQVMEGVPEMIPDIOVEATFPDGSKLVTNNMPI 99

RESULT 3
US-07-732-242C-1
Sequence 1, Application US/07732242C
Patent No. 5298399

GENERAL INFORMATION:
APPLICANT: Uozumi, Takeshi; Maseaki, Haruhiko;
APPLICANT: Hidaka, Makoto; Nakamura, Akira;
APPLICANT: Maeda, Michiisa; Yoneta, Yasuo
TITLE OF INVENTION: Gene of Urease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frlshauf, Holcz, Goodman & Woodward, P. C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72mb
COMPUTER: IBM PC compatible (NEC PC-9801 RX)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/732,242C
FILING DATE: 19910718
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPN 2-210178
FILING DATE: 10-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910532/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 972-1400
TELEFAX: (212) 370-1622
TELEX: 236268

INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-732-242C-1

Query Match 22.8%; Score 269.5; DB 1; Length 99;
Best Local Similarity 55.6%; Pred. No. 9e-23;

Matches 55; Conservative 18; Mismatches 25; Indels 1; Gaps 1;
Qy 2 KLTPEQEKFLYYAGVARRKAEGLKLNPEAIAYISAHIMDEARRGKKTVAQLMEEC 61
Db 1 KLTREMEKLMIVAAADLARRRERGLKLNPEAVAMITVELEGADG-KTVAQLMQYG 59
Qy 62 MHLFKDEVMPGVGNV/PDLGVEATFPDGTCLVTVNMP1 100
Db 60 ATLTKEQVMEGVAMIPDQIETFPDGTCLVTVHDP1 98

RESULT 4
US-08-467-822-23

; Sequence 23, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fimnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-822-23

Query Match 22.1%; Score 261.5; DB 2; Length 100;
Best Local Similarity 56.0%; Pred. No. 7.4e-22;
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;
Qy 1 VKLTPEQEKFLYYAGVARRKAEGLKLNPEAIAYISAHIMDEARRGKKTVAQLMEEC 60
Db 1 MELTPREKDLTLFTALVARRLAKGLKLNPERVALISCAIMEGARREG-KTVAQLMSE 59

Qy 61 CMHFLKDEVMPGVGNV/PDLGVEATFPDGTCLVTVNMP1 100
Db 60 GRTVLTAQVMEGVEMIKDQVECTFPDGTCLVSIHSP1 99

RESULT 5
US-08-432-697-23

; Sequence 23, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fimnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-432-697-23

Query Match 22.1%; Score 261.5; DB 3; Length 100;
Best Local Similarity 56.0%; Pred. No. 7.4e-22;
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;
Qy 1 VKLTPEQEKFLYYAGVARRKAEGLKLNPEAIAYISAHIMDEARRGKKTVAQLMEEC 60
Db 1 MELTPREKDLTLFTALVARRLAKGLKLNPERVALISCAIMEGARREG-KTVAQLMSE 59
Qy 61 CMHFLKDEVMPGVGNV/PDLGVEATFPDGTCLVTVNMP1 100
Db 60 GRTVLTAQVMEGVEMIKDQVECTFPDGTCLVSIHSP1 99

RESULT 6
US-08-466-248-23
; Sequence 23, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien

```

; APPLICANT: Ferrero, Richard L.
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-23

Query Match      22.1%; Score 261.5; DB 3; Length 100;
Best Local Similarity 56.0%; Pred. No. 7.4e-22;
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;

QY      1 VKITPEQEKFLYYAGEVARKKABGLKNOPEALAYISAHIMDEARRGKTYVAQLMEB 60
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1 MELTPREKDLTLFTAGLVAERLAKGLKLNPERVALISCAIMEGARREG-KTVAQLMSE 59

QY      61 CMHFLKKDEVMVPGVNMVPLDGEATFPDGTKLVTVNMP1 100
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      60 GRTVLTAEQVMEGVPEMINKDQVECTFPDGTKLVSTHSP1 99

RESULT 7
US-09-602-777A-18
; Sequence 18, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroege, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habethauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
```

```

; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 18
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-602-777A-18

Query Match      21.2%; Score 250.5; DB 4; Length 100;
Best Local Similarity 52.0%; Pred. No. 1.3e-20;
Matches 51; Conservative 20; Mismatches 26; Indels 1; Gaps 1;

QY      3 LTPKEQEKFLYYAGEVARKKABGLKNOPEALAYISAHIMDEARRGKTYVAQLMEB 62
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      3 ITPREQEKLMIVAADLARRRKDRGLKLNHPRAVALITYELLEGARDG-RTVADLMWGS 61

QY      63 HFLKKDEVMVPGVNMVPLDGEATFPDGTKLVTVNMP1 100
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      62 TLTTRDVLAEGLPEMIPDQVEATFPDGTKLVTVNMP1 99
```

RESULT 8
US-09-602-777A-12
Sequence 12, Application US/09602777A
Patent No. 6831165
GENERAL INFORMATION:
APPLICANT: Pompejun, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zeidler, Oskar
APPLICANT: Haberer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
FILE REFERENCE: BGI-128CP
CURRENT APPLICATION NUMBER: US/09/602,777A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932127.2
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932129.9
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: DE 19932226.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932920.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932924.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932928.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932930.3
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932933.8
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932935.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932973.7
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933002.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933003.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941390.8
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941391.6
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 442
SEQ ID NO 12
LENGTH: 40
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-602-777A-12
Query Match 5.8%, Score 68.5; DB 4; Length 40;

Best Local Similarity 31.1%; Pred. No. 1.5;
Matches 14; Conservative 8; Mismatches 16; Indels 7; Gaps 1;
Cy 102 PDEHFKAGEVFGGCDKDIENLAKGEVTELEVTNEGPKSLHYGSHF 146
Db 3 FGEYILSESLTG-----NVGREAKTIEIINTGDRPVQGSHF 40

RESULT 9
US-09-640-211A-973
Sequence 973, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Modification of Gene Transcription
FILE REFERENCE: 11000.1021CIU
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 973
LENGTH: 97
TYPE: PRT
ORGANISM: Pinus radiata
US-09-640-211A-973

Query Match 5.2%; Score 62; DB 4; Length 97;
Best Local Similarity 31.0%; Pred. No. 32;
Matches 18; Conservative 14; Mismatches 18; Indels 8; Gaps 3;

Cy 114 GCDKDIIE-LNAGKEVTELEVTNEGPKSLHYGSHFHFPEANKALKPDERAYGRLDI 169
Db 39 GCERKVNALNMSMGVSVVERQYKATVTG-----YVDANKVLK--RVKQTGKKAL 90

RESULT 10
US-09-248-796A-27427
Sequence 27427, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 27427
LENGTH: 63
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-27427

Query Match 5.1%; Score 60; DB 4; Length 63;
Best Local Similarity 30.9%; Pred. No. 28;
Matches 17; Conservative 9; Mismatches 25; Indels 4; Gaps 2;
Cy 168 DIPSGNTLRIGAGQTRKVLIPLGSKKVIQMGGLVNNIDERRHKALPKAKSH 222
Db 11 DIASVVGISGKNV--ADKLHLGSGVHTVG--GVAGGIAGNLEHAKADEVREH 61

RESULT 11
US-08-945-038-12
Sequence 12, Application US/08945038

```

Patent No. 6762295
GENERAL INFORMATION:
APPLICANT: Doidge, Christopher V.
APPLICANT: Lee, Adrian
APPLICANT: Radcliff, Fiona J.
APPLICANT: Hocking, Diana M.
APPLICANT: Webb, Elizabeth A.
TITLE OF INVENTION: PROTECTIVE HELICOBACTER ANTIGENS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,038
FILING DATE: 23-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00225
FILING DATE: 19-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN2575
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN3931
FILING DATE: 03-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7565
FILING DATE: 16-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 017227/0133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
US-08-945-038-12

Query Match 4.9%; Score 58; DB 4; Length 18;
Best Local Similarity 61.1%; Pred. No. 7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VKLTPKEOKFLYYAGE 18
:|||||:|:|
Db 1 MKLTPKEDKMLHRAGE 18

RESULT 12
US-09-270-767-60797
Sequence 60797, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17

```

```

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60797
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-60797

Query Match          4.8%; Score 57; DB 4; Length 83;
Best Local Similarity 28.6%; Pred. No. 92;
Matches 24; Conservative 13; Mismatches 37; Indels 10; Gaps 3;

QY      40 SAHIMDEARRGKKTVAQLMEBCNHFLLKKDEVMGVMVDPD---LGVETFPDGTLYVT    95
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1 AAOCTAASAGATTTISLMTKTVFHFAPGENFKADGYVTEHTERLLKEHLARTGKVHT    60

QY      96 VMNPIDPD-----EHFKAGEVKKG    114
       :|::||::||::||::||::||::||::||::||::||::||::||::||
Db      61 -RPPPEPNGILHGHAAYININFG    83


RESULT 13
US-09-513--999C-4321
Sequence 4321, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: S9.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 4321
LENGTH: 92
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -20...-1
OTHER INFORMATION: score 3.5
OTHER INFORMATION: seq IVIGTGLTECILS/GI
FEATURE:
NAME/KEY: UNSURE
LOCATION: 43
OTHER INFORMATION: Xaa=Pro or Thr
FEATURE:
NAME/KEY: UNSURE
LOCATION: 51
OTHER INFORMATION: Xaa=Asp or Glu
US-09-513--999C-4321

Query Match          4.8%; Score 57; DB 4; Length 92;
Best Local Similarity 30.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 3; Mismatches 22; Indels 18; Gaps 2;

QY      153 KALKFDEKKYG-----KRLDIPSGNTLRIGAGQTRKVLPIPGSKKVIQM    199
       ||::||::||::||::||::||::||::||::||::||::||::||::||
Db      29 KVLHMDSNPYYGGSSASITPLELDYKRFXKIPGSFXESMGRKXNVVDLP-----KFLMA    83

QY      200 NG 201
       ||
Db      84 NG 85

RESULT 14
US-09-328-352-7014
Sequence 7014, Application US/09328352
```

```

; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7014
; LENGTH: 100
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7014

```

```

Query Match          4.8%; Score 57; DB 4; Length 100;
Best Local Similarity 25.0%; Pred. No. 1.2e+02;
Matches 11; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

```

```

OY      111 VKPGCDKDIENAGKEVTELEVTNEGPKSLHVGSHFHFPEANKA 154
          |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      20 VELTCSYIHTPSGASGNKLGIDNRAEDLIHITDPSFPKTEA 63

```

```

RESULT 15
US-09-248-796A-23863
; Sequence 23863, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23863
; LENGTH: 83
; TYPE: PR1
; ORGANISM: Candida albicans
US-09-248-796A-23863

```

```

Query Match          4.8%; Score 56.5; DB 4; Length 83;
Best Local Similarity 31.8%; Pred. No. 1.1e+02;
Matches 14; Conservative 8; Mismatches 19; Indels 3; Gaps 1;

```

```

OY      122 NAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKA 162
          |||:|||||:|||||:|||||:|||||:|||||:|||||
Db      4  NQGEIERIKIVAKRPTCVTIGSGIKNNEMFVTGHLKLDDEKA 47

```

```

Search completed: February 15, 2005, 20:44:31
Job time : 9.53904 secs

```


This Page Blank (uspto)

LENGTH: 2735 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-920-095-1

Alignment Scores:

Pred. No.:	1-2e-70	length:	2735
Score:	610.50	Matches:	118
Percent Similarity:	71.88%	Conservative:	43
Best Local Similarity:	52.68%	Mismatches:	62
Query Match:	51.69%	Indels:	1
DB:	2	Gaps:	1

US-09-904-994B-2 (1-226) x US-08-920-095-1 (1-2735)

QY 1 VallylsleuThProlysgluingluysPheleuLeuTyTyxAlaglygluVala 20
DB 102 ATGAAACTACCCCAAAAGATTAGATTGATGCTCCTACCTGAGATTGGCT 161
QY 21 ArglysaArglysaIaglyleuysleuaerGlnProgluAllealatyrlIeser 40
DB 162 AAAAAACGCAAGAAAGCATTAACTTAATGTAGAACAGTAGCTTGTGATTAGT 221
QY 41 AlaHieIlleMerAspGluAaArgGlylylyleThValAlagInleuMetglu 60
DB 222 GCCCATTTATGAGAAAGCGAAGCTGTAAAGACCTGCGCTGATTGATGCAAGAA 281
QY 61 CysMetHisPheleuyslysaerGluValMetProglYValGlyAsnMetValProasp 80
DB 282 GGGCGCAGCTCTTTAAACAGATGATGTGATGATGGCGTGCAAGCATGATCCATGA 341
QY 81 LeuGlyValGluAlaThrPheProaspGlyThlyleuValThraGlnGlyProlySerLeu 100
DB 342 GTGGGTATTGAACGAGTGTCTCTGATGGACTAACTCGTAACCGTGATACCCCTATT 401
QY 101 GluProaspGluHisPheleuysIaglygluVallylePheGlyCysaspLyasapIleGlu 120
DB 402 GAGGCCAATGTAAATTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
QY 121 LeuAsnAlaGlylysgluValThrGluLeuGluValThraGlnGlyProlySerLeu 140
DB 459 ATCAACGAGGCAAAAGCCGTTAGCGTGAAGATTAAAGTTGTCGCGCAGACCGGTT 518
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnlysaIleuysPheAspArgGlu 160
DB 519 CAATGCGCTCACTTCATCTTCTTGAAGTGAATGATGCTTGAACAGAA 578
QY 161 lySAIATyTgLylysaArgleuAspIleProSerGlyAsnThrleuArgIleGlyAlaGly 180
DB 579 AAAAATTGCGTAAACCTTAAAGCATTTGCGACGCGGACACCGGTAAGATTGAGCTTGGC 638
QY 181 GlnThrArglysaValGlnleuIleProleuGlyGlySerlylysaValIleGlyMetAsn 200
DB 639 GAAGAAATCCCTGAATGATTGACATTTGCGGTAACAGAGAATCTTGTGATTAAAC 698
QY 201 GlyleuValAsnAsnIleAlaAspGluArgHislyleHislyleAlaLeuAspLySaIyS 220
DB 699 GCATTGTGATGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 758
QY 221 SerHisGlyPhe 224
DB 759 GAGCGTGTGTTT 770

RESULT 2

PCT-US96-05800-1

Sequence 1, Application PC/TUS9605800

GENERAL INFORMATION:

APPLICANT: Oryax, Inc.

TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05800

FILING DATE: 23-APR-1996

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/431,041

FILING DATE: 28-APR-1995

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/568,122

FILING DATE: 06-DEC-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 06132/020001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2735 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US96-05800-1

PCT-US96-05800-1 (1-2735)

Alignment Scores:	Pred. No.:	1-2e-70	length:	2735
Score:	610.50	Matches:	118	
Percent Similarity:	71.88%	Conservative:	43	
Best Local Similarity:	52.68%	Mismatches:	62	
Query Match:	51.69%	Indels:	1	
DB:	5	Gaps:	1	

US-09-904-994B-2 (1-226) x PCT-US96-05800-1 (1-2735)

QY 1 VallylsleuThProlysgluingluysPheleuLeuTyTyxAlaglygluVala 20
DB 102 ATGAAACTACCCCAAAAGATTAGATTGATGCTCCTACCTGAGATTGGCT 161
QY 21 ArglysaArglysaIaglyleuysleuaerGlnProgluAllealatyrlIeser 40
DB 162 AAAAAACGCAAGAAAGCATTAACTTAATGTAGAACAGTAGCTTGTGATTAGT 221
QY 41 AlaHieIlleMerAspGluAaArgGlylylyleThValAlagInleuMetglu 60
DB 222 GCCCATTTATGAGAAAGCGAAGCTGTAAAGACCTGCGCTGATTGATGCAAGAA 281
QY 61 CysMetHisPheleuyslysaerGluValMetProglYValGlyAsnMetValProasp 80
DB 282 GGGCGCAGCTCTTTAAACAGATGATGTGATGATGGCGTGCAAGCATGATCCATGA 341
QY 81 LeuGlyValGluAlaThrPheProaspGlyThlyleuValThraGlnGlyProlySerLeu 100
DB 342 GTGGGTATTGAACGAGTGTCTCTGATGGACTAACTCGTAACCGTGATACCCCTATT 401
QY 101 GluProaspGluHisPheleuysIaglygluVallylePheGlyCysaspLyasapIleGlu 120
DB 402 GAGGCCAATGTAAATTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458


```

: LENGTH: 2619 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: 31..36
: OTHER INFORMATION: /standard_name="Shine-Dalgarno
: OTHER INFORMATION: sequence."
:
: NAME/KEY: misc feature
: LOCATION: 756..759
: OTHER INFORMATION: /standard_name="Shine-Dalgarno
: OTHER INFORMATION: sequence."
:
US-08-467-822-19

Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
DB: Gaps: 1

US-09-904-994b-2 (1-226) x US-08-467-822-19 (1-2619)
QY 1 VallyleuthrProlysgIngluLyphleuLeuTyTyraIagIygluVala 20
DB 43 ATGAACTAAGCGCTAAAGAACTAGACAGTAATGCTCATTCGCGGAGATGGCA 102
QY 21 ArgLyArgLyAgluLyLeuLyLeuAngInProgluAaIaIaIaTyIIseser 40
DB 103 GAAGAACGCTGGCGCTGTGTAACATTAACAAGCGGCTGCATTAGC 162
QY 41 AlAhIeIlemerAsgIuaIaArgArgLylyeIythrValaIagInIlemerGlu 60
DB 163 GGGCGTGTGATGAAAGGCGCGTGAATTAAGCGTGGCGATTGATGCAAGAA 222
QY 61 CyMetHisPheLeuLySAspGluValMetProgluValGlyAsnMetValProasp 80
DB 223 GCGAGACCTTGCTTAAAGAAAGAAATGTATGACGCGCTAGACATGATTCAGAA 282
QY 81 LeuGlValGluValThrPheProAspGlyThryIyIeValuIThrValaenITpProIle 100
DB 283 GTGGGATTAAGCTAACTTCCCGAGTGAACCAAGCTTGAATCACTCCGCTA 342
QY 101 GluProAspGluHisPheLySAlaGlyValLyPheGlyCyAspLyAspIleGlu 120
DB 343 GAGGATTAATGCAAAATTAGCCCCGCGAGGTC---TTCTTAAATAATGAGACATTACT 399
QY 121 LeuAsnAaGlyLySgluValIThrGluLeuGluValIThrAsnGluGlyProLySserLeu 140
DB 400 ATTAACGCGCGCAAGAAAGCATTAGCTTGAAGTGAATAATTAAGCGCATTCGCTGTG 459
QY 141 HisValGlySeriHisPheHisPhePheGluAlaAsnLySAlaLeuLySPhaAspArgGlu 160
DB 460 CAGGTGGATCAACTTTCCTTCTTGAAGTGAATTAAGCTTGAAGCTTGAAGCTTGA 519
QY 161 LyAlaTyGlyLySArgLyLeuAspIleProSerGlyAsnThrIleuAaGlyIleGlyAlaGly 180
DB 520 AAAAGCTTTTGCAAAAGCCCTTACATTCGATTCGGAACAGCGGCTTGAACCGGG 579
QY 181 GluThrArgLyValGluLeuIleProLeuGlyGlySeriLySlyValIleGlyMetAsn 200
DB 580 GAGGAAAAAGTGTGAACTCATTCGATCGCGGGAATTAAGCGCATCTATGCTTAAT 639
QY 201 GlyLeuValAsnAsnIleAlaAspGluArgHisIyHisIySAlaLeuAspLySAlaLyS 220
DB 640 TCTTGTGGATGCGCAAGCGCATGCGATGTAAATAACTCGGCTTAAACGCGCTAA 699
QY 221 SerHisGlyPhe 224
|||||
```

```

DB 700 GAAAGGTTTT 711

RESULT 6
US-08-432-697-19
: Sequence 19, Application US/08432697
: Patent No. 6248330
: GENERAL INFORMATION:
: APPLICANT: Labigne, Agnes
: APPLICANT: Sauerbaum, Sebastien
: APPLICANT: Ferreira, Richard L.
: APPLICANT: Thibierge, Jean-Michel
: TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
: TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
: TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSER: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D. C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/432,697
: FILING DATE: 02-MAY-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 03495.0137-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2619 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: 31..36
: OTHER INFORMATION: /standard_name="Shine-Dalgarno
: OTHER INFORMATION: sequence."
:
: NAME/KEY: misc feature
: LOCATION: 756..759
: OTHER INFORMATION: /standard_name="Shine-Dalgarno
: OTHER INFORMATION: sequence."
:
US-08-432-697-19

Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
DB: Gaps: 1

US-09-904-994b-2 (1-226) x US-08-432-697-19 (1-2619)
QY 1 VallyleuthrProlysgIngluLyphleuLeuTyTyraIagIygluVala 20
DB 43 ATGAACTAAGCGCTAAAGAACTAGACAGTAATGCTCATTCGCGGAGATGGCA 102
```

```
QY 21 ArglySArGlyAlaGluGlyLeuAenGlnProGluAlaIleAlaTyrlIeser 40
DB 103 GAAGAAGCGTTGGCGGCGTGTGTGAACCTTCACTTACACGAAGCGGTCCGCTCATTAAC 162
QY 41 AlaHsIleMeLaspGluAlaArgArgGlyLysLysThrValAlaGluLeuMetGluGlu 60
DB 163 GCGCGGTGATGAGAAAGCGCGGTGATGATTAATAAAGCGGTGCGGATTTGATGCAAGAA 222
QY 61 CysMetHisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAsp 80
DB 223 GCGAGAGACTTGGCTTAAATAAAGAAATGATGAGACGCGGTGACCAAGATGATCAATGAA 282
QY 81 LeuGlyValAlaGluAlaThrPheProAspGlyThrLysLeuValThrValAsnTrpProIle 100
DB 283 GTGGGATTTGAAGCTTAACCTTCCCGCATGGAACCAACTTGTAACTATCCACACTCCGGTA 342
QY 101 GluProAspGluHisPheLysAlaGlyLysValLysPheGlyCysAspLysAspIleGlu 120
DB 343 GAGGATTAATGCAAAATTACCCCGCGCAGAGTC---TTCTTAAATAAATGAGACATTACT 399
QY 121 LeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLysSerLeu 140
DB 400 ATTTAAGCGCGGCAAAAGCCACTTACCTTGAAGTAAATAAAGCGCATGCTCTGTG 459
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
DB 460 CAGGTGGGATTCACATTTCACCTTCTCGAAGTGAATTAAGCTTGTGACCTTCATTCGCGCA 519
QY 161 LysAlaTyrglyLysArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
DB 520 AAAAGCTTTGCAAAAGCGCTAGCATGTGATGATGCAAGACGCGTGGCTTGAACCGCGG 579
QY 181 GlnThrArgLysValGlnLeuLysProLeuGlyLysSerLysLysValIleGlyMetAsn 200
DB 580 GAGGAAAAAGTGTGACTTGAACATCGCGGGAATAAGGCAATCTTAATGGCTTTAAT 639
QY 201 GlyLeuValAsnAsnIleAlaAspGluAArgHisLysHisLysAlaLeuAspLysAlaLys 220
DB 640 TCTTGTGTGATGCCCAAGCCGATGCCGATGGTAAATAAAGCTTAAAGCGCTTAA 699
QY 221 SerHisGlyPhe 224
DB 700 GAAAAAGGTTTT 711

RESULT 7
US-08-466-248-19
; Sequence 19, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerdam, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOCTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
```

```
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495, 0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-466-248-19

Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
DB: Gaps: 1

US-09-904-994B-2 (1-226) x US-08-466-248-19 (1-2619)
QY 1 ValLysLeuThrProLysGluGlnGlyLysPheLeuLeuTyrrAlaGlyLysValAla 20
DB 43 ATGAAGCTTAAGCCCTTAAGAACTAGACAAATTAAAGCTTCATTATGCGGCGAGATTGGCA 102
QY 21 ArglySArGlyAlaGluGlyLeuAenGlnProGluAlaIleAlaTyrlIeser 40
DB 103 GAAGAAGCGTTGGCGGCGTGTGTGAACCTTCACTTACACGAAGCGGTCCGCTCATTAAC 162
QY 41 AlaHsIleMeLaspGluAlaArgArgGlyLysLysThrValAlaGluLeuMetGluGlu 60
DB 163 GCGCGGTGATGAGAAAGCGCGGTGATGATTAATAAAGCGGTGCGGATTTGATGCAAGAA 222
QY 61 CysMetHisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAsp 80
DB 223 GCGAGAGACTTGGCTTAAATAAAGAAATGATGAGACGCGGTGACCAAGATGATCAATGAA 282
QY 81 LeuGlyValAlaGluAlaThrPheProAspGlyThrLysLeuValThrValAsnTrpProIle 100
DB 283 GTGGGATTTGAAGCTTAACCTTCCCGCATGGAACCAACTTGTAACTATCCACACTCCGGTA 342
QY 101 GluProAspGluHisPheLysAlaGlyLysValLysPheGlyCysAspLysAspIleGlu 120
DB 343 GAGGATTAATGCAAAATTACCCCGCGCAGAGTC---TTCTTAAATAAATGAGACATTACT 399
QY 121 LeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLysSerLeu 140
DB 400 ATTTAAGCGCGGCAAAAGCCACTTACCTTGAAGTGAATAAAGCGCATGCTCTGTG 459
```


TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-08-956-171E-22

Alignment Scores:
Pred. No.: 8,4e-53 Length: 5966
Score: 480.50 Matches: 101
Percent Similarity: 63.35% Conservative: 39
Best Local Similarity: 45.70% Mismatches: 69
Query Match: 40.69% Indels: 12
DB: 4 Gaps: 3

US-09-904-994b-2 (1-226) x US-08-956-171E-22 (1-5966)

```

QY      4  ThrProLysGluGlnGluLysPheLeuLeuTyrrYrAlaGluValAlaArgLysArg 23
DB      4003  ACACAAAGAGAGACAAATTAATGATGATGCGCGCGAGAGTTGACGTCGTCGT 3944
QY      24  LysAlaGluGluLysLysLeuAnglProGluAlaIleAlaTyrlIleSerAlaHisIle 43
DB      3943  AAGGACGCTGGTTGAACTAAATCATCTCAGAGCACTTACCTTTAATCAGCGATGAATTA 3884
QY      44  MetAspGluAlaArgArgLysLysLysThrValAlaGlnLeuMetGluGluCysMetHis 63
DB      3883  TTGAAAGGTGCACGCGATGCT--AAGACCGTTGCAGAGATTATGATTATGTTAGACAA 3827
QY      64  PheLeuLysLysAspGluValMetProGluValGlyAsnMetValProAspLeuGlyVal 83
DB      3826  ATTCTTAACAAGAAAGATGTCATGATGCTGCAACACATGATTAACAGATATCGAAATC 3767
QY      84  GluAlaThrPheProAspGlyThrLysLeuValThrValAsnProPro-IleGluProAs 103
DB      3766  GAGGCTACGTTCCCGCATGCTACTAATGATTATACACGATATACCTTATG----- 3715
QY      103  pGluHisPheLysAlaGlyLysValLysPheGlyCysAsp-----LysAs 118
DB      3714  -----TTTAAAGAGAGCGATCAATGATACAGAGAAATTATTAACAAAAGTACAG 3662
QY      118  pIleGluLeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLys 138
DB      3661  GGTTGAATTAATTAACCATCATCTGAAACAGATTGAAAGTTGAAATAACAGAAATCG 3602
QY      138  sSerLeuHisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAs 158
DB      3601  ACCAATTCAGATGGGCTCACATTTTCATTTTATGAAGCAAAATGACGACATTGATTCGA 3542
QY      158  pArgGluLysValArgLysLysLysArgLeuAspIleProSerGlyAsnThrLeuArgIleG 178
DB      3541  ACGTGAATGCGCATATGAAAAACATTTAGATATTCAGAGTGGAGCGAGCTGTCGATTTGA 3482
QY      178  ValaGluGlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysLeuValIleG 198
DB      3481  ACCTGGGGAATAAAAGAGTTCAATTAAGTATGCTGGCAAAACGTAAATTTTGG 3422
QY      198  YMetAsnGlyLeuValAsnAsnIleAlaAspGluArgHisLysHisLysAlaLeuAspLys 218
DB      3421  TTTTCGTGATGTCATATGCTCTATCGATGAGTCAACGTCGTATGATGCCCACTGATGA 3362
QY      218  g 218
DB      3361  A 3361

```

RESULT 10
US-08-781-986A-22/c
Sequence 22, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:

```

/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,986A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5966 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ US-08-781-986A-22

```

Alignment Scores:
Pred. No.: 8,4e-53 Length: 5966
Score: 480.50 Matches: 101
Percent Similarity: 63.35% Conservative: 39
Best Local Similarity: 45.70% Mismatches: 69
Query Match: 40.69% Indels: 12
DB: 4 Gaps: 3

US-09-904-994b-2 (1-226) x US-08-781-986A-22 (1-5966)

```

QY      4  ThrProLysGluGlnGluLysPheLeuLeuTyrrYrAlaGluValAlaArgLysArg 23
DB      4003  ACACAAAGAGACAAACAAATTAATGATGATGCGCGCGAGAGTTGACGTCGTCGT 3944
QY      24  LysAlaGluGluLysLysLeuAnglProGluAlaIleAlaTyrlIleSerAlaHisIle 43
DB      3943  AAGGACGCTGGTTGAACTAAATCATCTCAGAGCACTTACCTTTAATCAGCGATGAATTA 3884
QY      44  MetAspGluAlaArgArgLysLysLysThrValAlaGlnLeuMetGluGluCysMetHis 63
DB      3883  TTGAAAGGTGCACGCGATGCT--AAGACCGTTGCAGAGATTATGATTATGTTAGACAA 3827
QY      64  PheLeuLysLysAspGluValMetProGluValGlyAsnMetValProAspLeuGlyVal 83
DB      3826  ATTCTTAACAAGAAAGATGTCATGATGCTGCAACACATGATTAACAGATATCGAAATC 3767
QY      84  GluAlaThrPheProAspGlyThrLysLeuValThrValAsnProPro-IleGluProAs 103
DB      3766  GAGGCTACGTTCCCGCATGCTACTAATGATTATACAGATATACACCTTATG----- 3715
QY      103  pGluHisPheLysAlaGlyLysValLysPheGlyCysAsp-----LysAs 118
DB      3714  -----TTTAAAGAGAGCGATCAATGATACAGAGAAATTATTAACAAAAGTACAG 3662
QY      118  pIleGluLeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLys 138
DB      3661  GGTTGAATTAATTAACCATCATCTGAAACAGATTGAAAGTTGAAATAACAGAAATCG 3602
QY      138  sSerLeuHisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAs 158
DB      3601  ACCAATTCAGATGGGCTCACATTTTCATTTTATGAAGCAAAATGACGACATTGATTCGA 3542

```



```

1 STREET: 2190 Commons Parkway
2 City: Okemos
3 STATE: Michigan
4 COUNTRY: USA
5 ZIP: 48864
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
9 MEDIUM TYPE: storage
10 COMPUTER: IBM Compatible
11 OPERATING SYSTEM: MS-DOS
12 SOFTWARE: Wordperfect 5.1
13
14 CURRENT APPLICATION DATA:
15 FILING DATE: July 26, 1996
16 CLASSIFICATION: 435
17
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER:
20 FILING DATE:
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Ian C. McLeod
23
24 REGISTRATION NUMBER: 20,931
25 REFERENCE/DOCKET NUMBER: MSU 4.1-309
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (517) 347-4100
28 TELEFAX: (517) 347-4103
29
30 TELEX: No. 5846752e
31
32 INFORMATION FOR SEQ ID NO: 1:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 2400
35 TYPE: nucleotides
36 STRANDEDNESS: Single
37 TOPOLOGY: Linear
38
39 MOLECULE TYPE:
40 DESCRIPTION: cDNA
41 HYPOTHETICAL: NO
42 ANTI-SENSE: NO
43 ORIGINAL SOURCE:
44 ORGANISM: Klebsiella aerogenes
45 STRAIN: CG253
46 INDIVIDUAL ISOLATE:
47 CELL TYPE: N/A
48
49 FEATURE:
50 NAME/KEY: cDNA encoding mutant urease
51 NAME/KEY: 'H2190
52 LOCATION: Modification at position 1312 to
53 LOCATION: glutamine
54 IDENTIFICATION METHOD: Sequencing
55 OTHER INFORMATION:
56 ?
57 ?
58 ?
59 ?
60 ?
61 ?
62 ?
63 ?
64 ?
65 ?
66 ?
67 ?
68 ?
69 ?
70 ?
71 ?
72 ?
73 ?
74 ?
75 ?
76 ?
77 ?
78 ?
79 ?
80 ?
81 ?
82 ?
83 ?
84 ?
85 ?
86 ?
87 ?
88 ?
89 ?
90 ?
91 ?
92 ?
93 ?
94 ?
95 ?
96 ?
97 ?
98 ?
99 ?
100 ?
101 ?
102 ?
103 ?
104 ?
105 ?
106 ?
107 ?
108 ?
109 ?
110 ?
111 ?
112 ?
113 ?
114 ?
115 ?
116 ?
117 ?
118 ?
119 ?
120 ?
121 ?
122 ?
123 ?
124 ?
125 ?
126 ?
127 ?
128 ?
129 ?
130 ?
131 ?
132 ?
133 ?
134 ?
135 ?
136 ?
137 ?
138 ?
139 ?
140 ?
141 ?
142 ?
143 ?
144 ?
145 ?
146 ?
147 ?
148 ?
149 ?
150 ?
151 ?
152 ?
153 ?
154 ?
155 ?
156 ?
157 ?
158 ?
159 ?
160 ?
161 ?
162 ?
163 ?
164 ?
165 ?
166 ?
167 ?
168 ?
169 ?
170 ?
171 ?
172 ?
173 ?
174 ?
175 ?
176 ?
177 ?
178 ?
179 ?
180 ?
181 ?
182 ?
183 ?
184 ?
185 ?
186 ?
187 ?
188 ?
189 ?
190 ?
191 ?
192 ?
193 ?
194 ?
195 ?
196 ?
197 ?
198 ?
199 ?
200 ?
201 ?
202 ?
203 ?
204 ?
205 ?
206 ?
207 ?
208 ?
209 ?
210 ?
211 ?
212 ?
213 ?
214 ?
215 ?
216 ?
217 ?
218 ?
219 ?
220 ?
221 ?
222 ?
223 ?
224 ?
225 ?
226 ?
227 ?
228 ?
229 ?
230 ?
231 ?
232 ?
233 ?
234 ?
235 ?
236 ?
237 ?
238 ?
239 ?
240 ?
241 ?
242 ?
243 ?
244 ?
245 ?
246 ?
247 ?
248 ?
249 ?
250 ?
251 ?
252 ?
253 ?
254 ?
255 ?
256 ?
257 ?
258 ?
259 ?
260 ?
261 ?
262 ?
263 ?
264 ?
265 ?
266 ?
267 ?
268 ?
269 ?
270 ?
271 ?
272 ?
273 ?
274 ?
275 ?
276 ?
277 ?
278 ?
279 ?
280 ?
281 ?
282 ?
283 ?
284 ?
285 ?
286 ?
287 ?
288 ?
289 ?
290 ?
291 ?
292 ?
293 ?
294 ?
295 ?
296 ?
297 ?
298 ?
299 ?
300 ?
301 ?
302 ?
303 ?
304 ?
305 ?
306 ?
307 ?
308 ?
309 ?
310 ?
311 ?
312 ?
313 ?
314 ?
315 ?
316 ?
317 ?
318 ?
319 ?
320 ?
321 ?
322 ?
323 ?
324 ?
325 ?
326 ?
327 ?
328 ?
329 ?
330 ?
331 ?
332 ?
333 ?
334 ?
335 ?
336 ?
337 ?
338 ?
339 ?
340 ?
341 ?
342 ?
343 ?
344 ?
345 ?
346 ?
347 ?
348 ?
349 ?
350 ?
351 ?
352 ?
353 ?
354 ?
355 ?
356 ?
357 ?
358 ?
359 ?
360 ?
361 ?
362 ?
363 ?
364 ?
365 ?
366 ?
367 ?
368 ?
369 ?
370 ?
371 ?
372 ?
373 ?
374 ?
375 ?
376 ?
377 ?
378 ?
379 ?
380 ?
381 ?
382 ?
383 ?
384 ?
385 ?
386 ?
387 ?
388 ?
389 ?
390 ?
391 ?
392 ?
393 ?
394 ?
395 ?
396 ?
397 ?
398 ?
399 ?
400 ?
401 ?
402 ?
403 ?
404 ?
405 ?
406 ?
407 ?
408 ?
409 ?
410 ?
411 ?
412 ?
413 ?
414 ?
415 ?
416 ?
417 ?
418 ?
419 ?
420 ?
421 ?
422 ?
423 ?
424 ?
425 ?
426 ?
427 ?
428 ?
429 ?
430 ?
431 ?
432 ?
433 ?
434 ?
435 ?
436 ?
437 ?
438 ?
439 ?
440 ?
441 ?
442 ?
443 ?
444 ?
445 ?
446 ?
447 ?
448 ?
449 ?
450 ?
451 ?
452 ?
453 ?
454 ?
455 ?
456 ?
457 ?
458 ?
459 ?
460 ?
461 ?
462 ?
463 ?
464 ?
465 ?
466 ?
467 ?
468 ?
469 ?
470 ?
471 ?
472 ?
473 ?
474 ?
475 ?
476 ?
477 ?
478 ?
479 ?
480 ?
481 ?
482 ?
483 ?
484 ?
485 ?
486 ?
487 ?
488 ?
489 ?
490 ?
491 ?
492 ?
493 ?
494 ?
495 ?
496 ?
497 ?
498 ?
499 ?
500 ?
501 ?
502 ?
503 ?
504 ?
505 ?
506 ?
507 ?
508 ?
509 ?
510 ?
511 ?
512 ?
513 ?
514 ?
515 ?
516 ?
517 ?
518 ?
519 ?
520 ?
521 ?
522 ?
523 ?
524 ?
525 ?
526 ?
527 ?
528 ?
529 ?
530 ?
531 ?
532 ?
533 ?
534 ?
535 ?
536 ?
537 ?
538 ?
539 ?
540 ?
541 ?
542 ?
543 ?
544 ?
545 ?
546 ?
547 ?
548 ?
549 ?
550 ?
551 ?
552 ?
553 ?
554 ?
555 ?
556 ?
557 ?
558 ?
559 ?
560 ?
561 ?
562 ?
563 ?
564 ?
565 ?
566 ?
567 ?
568 ?
569 ?
570 ?
571 ?
572 ?
573 ?
574 ?
575 ?
576 ?
577 ?
578 ?
579 ?
580 ?
581 ?
582 ?
583 ?
584 ?
585 ?
586 ?
587 ?
588 ?
589 ?
590 ?
591 ?
592 ?
593 ?
594 ?
595 ?
596 ?
597 ?
598 ?
599 ?
600 ?
601 ?
602 ?
603 ?
604 ?
605 ?
606 ?
607 ?
608 ?
609 ?
610 ?
611 ?
612 ?
613 ?
614 ?
615 ?
616 ?
617 ?
618 ?
619 ?
620 ?
621 ?
622 ?
623 ?
624 ?
625 ?
626 ?
627 ?
628 ?
629 ?
630 ?
631 ?
632 ?
633 ?
634 ?
635 ?
636 ?
637 ?
638 ?
639 ?
640 ?
641 ?
642 ?
643 ?
644 ?
645 ?
646 ?
647 ?
648 ?
649 ?
650 ?
651 ?
652 ?
653 ?
654 ?
655 ?
656 ?
657 ?
658 ?
659 ?
660 ?
661 ?
662 ?
663 ?
664 ?
665 ?
666 ?
667 ?
668 ?
669 ?
670 ?
671 ?
672 ?
673 ?
674 ?
675 ?
676 ?
677 ?
678 ?
679 ?
680 ?
681 ?
682 ?
683 ?
684 ?
685 ?
686 ?
687 ?
688 ?
689 ?
690 ?
691 ?
692 ?
693 ?
694 ?
695 ?
696 ?
697 ?
698 ?
699 ?
700 ?
701 ?
702 ?
703 ?
704 ?
705 ?
706 ?
707 ?
708 ?
709 ?
710 ?
711 ?
712 ?
713 ?
714 ?
715 ?
716 ?
717 ?
718 ?
719 ?
720 ?
721 ?
722 ?
723 ?
724 ?
725 ?
726 ?
727 ?
728 ?
729 ?
730 ?
731 ?
732 ?
733 ?
734 ?
735 ?
736 ?
737 ?
738 ?
739 ?
740 ?
741 ?
742 ?
743 ?
744 ?
745 ?
746 ?
747 ?
748 ?
749 ?
750 ?
751 ?
752 ?
753 ?
754 ?
755 ?
756 ?
757 ?
758 ?
759 ?
760 ?
761 ?
762 ?
763 ?
764 ?
765 ?
766 ?
767 ?
768 ?
769 ?
770 ?
771 ?
772 ?
773 ?
774 ?
775 ?
776 ?
777 ?
778 ?
779 ?
780 ?
781 ?
782 ?
```

```

Db      208 GGCCTGACGCTCTGTGACCCGCGACGAGGTGAGTGGGCGCTCCGGAATATATCCCGAT 267
Oy      81 LeuGlyValGluAlaThrPheProAspGlyThrIysLeuValThrValAsnTrpProIle 100
Db      268 ATCCAGGTGGAAGCCACTTCCCGGACGGCTCGAAGCTGTGACCGCTTACACCCGATTT 327
Oy      101 -----GluProAspGlu---HisPheIysAlaGlyGluValIysPhe 113
Db      328 ATCTGAGGTAGCGCCCATGATCCCGGTGAATATCATCGTTAAGCCCGGTCA----- 378
Oy      114 GlyCysAspIysAspIleGlyLeuAsnAlaGlyIysGluValThrGluLeuGluValThr 133
Db      379 -----ATAGCCCTGAATACCGCGCCGCAACCTGTCCGCTGCTGTTGAG 423
Oy      134 AsnGlyIProIysSerLeuHisValGlySerHisPheHisPheGluAlaAsnLys 153
Db      424 AACCAACGCGCATGCGCCGATTCAGGTGCGTTCCGACTACATTCGCGCGAGTTAACCCG 483
Oy      154 AlaLeuLysPheAspArgGlyLysAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 173
Db      484 GCGCTGAAGTTCACCGTACGACGAGCGCCGCGCTATCGCTGAATATATCCCGCGGACG 543
Oy      174 ThrLeuArgIleGlyAlaGlyIlnThrArgLysValGlnLeuIleProLeuGlyIysSer 193
Db      544 GCGGTACGCTTTGAACCCGCGCCAGAAAGCGACGAGCTGTGTGCTTCCCGGTAC 603
Oy      194 LysLysValIleGlyMetAsnGlyLeuVal 203
Db      604 CGCGCGCTCTTCGCGCTTCCGCGGAGGTCT 633

RESULT 13
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeehi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Dtease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frieauf, Holtz, Goodman & Woodward, P. C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910533/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6131 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both

```



```

Db      564511 ACAGGAAACAAAGATATGATCCAGCGCAATACCAATTA---GCCAAGCGCATATTC 564455
               |||||
Qy      121 LeuAenAlAGlyLyseGluValThrGluLeuGluValThrAsnGluGlyProlySerLeu 140
               |||||
Db      564454 GCTAATGTCGGCGAAGAAACCGTAAATCGAAGTAACAAATTCAGCGCACGCCCAATT 564395
               |||||
Qy      141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
               |||||
Db      564394 CAAATTGGCTCGCATTAACCAATTTTGGAAACCAATATATGCTTAATTTGACCGCACT 564335
               |||||
Qy      161 LysAlaTyGlyLyseArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
               |||||
Db      564334 TTGGCACCGTGAATGCGCTTAATGTTCCATCTGCGCAATGCGGTGCTTTGAACCGGT 564275
               |||||
Qy      181 GlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200
               |||||
Db      564274 GAAGTGAATCGAGTAATGATTGCTTTGGTGTATACCAATCATTTATGTTCCAT 564215
               |||||
Qy      201 GlyLeuValAsn 204
               |||||
Db      564214 AATCAATTGAT 564203
               |||||

RESULT 15
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6538289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Flaischmann
;             Mark D. Adams
;             Owen White
;             Hamilton O. Smith
;             J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
;                   the Haemophilus influenzae Rd Genome, Fragments
;                   Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

```

Alignment Scores:
Pred. No.:      3,79e-47      Length:      1830121
Score:          467.00      Matches:      96
Percent Similarity: 61.61%      Conservative: 42
Beet Local Similarity: 42.86%      Mismatches: 62
Query Match:      39.54%      Indels:      24
DB:              4          Gaps:      3

US-09-904-994B-2 (1-226) x US-09-643-990A-1 (1-1830121)
Qy      3 LeuThrProlyGluGlnGluLysPheLeuLeuTyTYrAlaGlyGluValAlaArgLys 22
               |||||
Db      564868 TTAACCTCCAGAGAAACAAGAAACAGATGCTTTCTCTCGCGGCGAACTTCGCGGCAAAA 564809
               |||||
Qy      23 ArgLysAlaGluGlyLeuLysLeuAsnGlnProGluAlaIleAlaTyTlSerAlaHis 42
               |||||
Db      564808 CCCAAAGCACCGCGGGCTTAAATTAACATATCCAGAAACTATTTGCTTAATTTCTAGTCAT 564749
               |||||
Qy      43 IleMetAspGluAlaArgArgGlyLysThrValAlaGlnLeuMetGluGlyMet 62
               |||||
Db      564748 TTACAAAGGCGCAGCAAGAAAGAA---ATGAGTGTAGCGGAAGTCAATGCAATATGCGCA 564692
               |||||
Qy      63 HisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAspLeuGly 82
               |||||
Db      564691 ACACTTTTAACCGTGTATGATGATCATGAGAGGTGTGGCGGAATGTTTCATGAAATGCCAG 564632
               |||||
Qy      83 ValGluAlaThrPheProAspGlyThrLysLeuValThrValAsnThrProIleGlu--- 101
               |||||
Db      564631 ATTGAAGCTACTTTCCCGCATGGACGAAACTTGTATCCGTCAATATCAATCAGATTA 564572
               |||||
Qy      101 -----
Db      564571 CCGAAGGTGGGCTTTAGCCCAAAATTAATATCAATGTTGGGCTTAAGCCCACTT 564512
               |||||
Qy      102 ---ProAspGlnHisPheLysAlaGlyGluValLysPheGlyCysAspLysAspIleGlu 120
               |||||
Db      564511 ACAGGAAACAAAGATATGATCCAGCGCAATACCAATTA---GCCAAGCGCATATTC 564455
               |||||
Qy      121 LeuAenAlAGlyLyseGluValThrGluLeuGluValThrAsnGluGlyProlySerLeu 140
               |||||
Db      564454 GCTAATGTCGGCGAAGAAACCGTAAATCGAAGTAACAAATTCAGCGCACGCCCAATT 564395
               |||||
Qy      141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
               |||||
Db      564394 CAAATTGGCTCGCATTAACCAATTTTGGAAACCAATATGCTTAATTTGACCGCACT 564335
               |||||
Qy      161 LysAlaTyGlyLyseArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
               |||||
Db      564334 TTGGCACCGTGAATGCGCTTAATGTTCCATCTGCGCAATGCGGTGCTTTGAACCGGT 564275
               |||||
Qy      181 GlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200
               |||||
Db      564274 GAAGTGAATCGAGTAATGATTGCTTTGGTGTATACCAATCATTTATGTTCCAT 564215
               |||||
Qy      201 GlyLeuValAsn 204
               |||||
Db      564214 AATCAATTGAT 564203
               |||||

```

Search completed: February 16, 2005, 00:59:42
Job time : 905.667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_nzp model

Run on: February 15, 2005, 20:03:24 ; Search time 31.5 Seconds
(without alignments)
13664.333 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 5147
Sequence: 1 rgragatcttcacacactc.....aaaaagtagaagccacagg 2883

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 648760

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_nzp.model -DEV=xlh
-Q=/cg2_1/USPTO.spool/US09904994/runat_11022005_134936_22927/app_query.fasta_1.3079
-DB=Issued_Patents_AA -OPMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62 -TRANS=numan40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptco -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=100
-USRR=US09904994 -CGN_1_1_26@runat_11022005_134936_22927 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA.*
1: /cg2_6/ptodata/1/iaa/5A.COMB.dep.*
2: /cg2_6/ptodata/1/iaa/5B.COMB.dep.*
3: /cg2_6/ptodata/1/iaa/6A.COMB.dep.*
4: /cg2_6/ptodata/1/iaa/6B.COMB.dep.*
5: /cg2_6/ptodata/1/iaa/PCTUS.COMB.dep.*
6: /cg2_6/ptodata/1/iaa/backfile1.dep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	281.5	5.5	100	1	US-08-967-513-3
2	281.5	5.5	100	1	US-08-967-513-3
3	269.5	5.2	99	1	US-07-732-242C-1
4	261.5	5.1	100	2	US-08-467-822-23
5	261.5	5.1	100	3	US-08-432-697-23
6	261.5	5.1	100	3	US-08-466-248-23
7	250.5	4.9	100	4	US-09-602-777A-18
8	87.5	1.7	92	4	US-09-252-991A-17446
9	78.5	1.5	96	4	US-09-328-352-4313
10	77	1.5	15	4	US-09-928-081-1
11	77	1.5	15	4	US-09-338-920B-1
12	71.5	1.4	93	4	US-09-489-039A-7356

C 13	70	1.4	89	4	US-09-489-039A-12832	Sequence 12832, A
C 14	69	1.3	99	4	US-09-328-352-4788	Sequence 4788, Ap
C 15	68.5	1.3	40	4	US-09-602-777A-12	Sequence 12, Appl
C 16	68	1.3	89	4	US-09-252-991A-17666	Sequence 17666, A
C 17	67.5	1.3	95	4	US-09-513-999C-7491	Sequence 7491, Ap
C 18	67	1.3	61	4	US-09-270-767-40422	Sequence 40422, A
C 19	67	1.3	61	4	US-09-270-767-55638	Sequence 55638, A
C 20	67	1.3	84	4	US-09-270-767-57094	Sequence 57094, A
C 21	67	1.3	96	4	US-09-513-999C-6065	Sequence 6065, Ap
C 22	67	1.3	99	2	US-08-537-400-19	Sequence 19, Appl
C 23	66.5	1.3	89	4	US-09-107-532A-3751	Sequence 3751, Ap
C 24	66.5	1.3	97	4	US-09-252-991A-27707	Sequence 27707, A
C 25	66	1.3	96	4	US-09-270-767-16576	Sequence 36576, A
C 26	66	1.3	96	4	US-09-270-767-51793	Sequence 51793, A
C 27	65.5	1.3	62	4	US-09-621-976-4501	Sequence 4501, Ap
C 28	65	1.3	75	4	US-09-471-780C-126	Sequence 5840, Ap
C 29	65	1.3	84	4	US-09-471-276-1347	Sequence 1347, Ap
C 30	64.5	1.3	81	5	PCT-US95-06266-126	Sequence 126, App
C 31	64	1.2	15	3	US-09-091-001-2	Sequence 2, Appl
C 32	64	1.2	95	4	US-09-252-991A-25355	Sequence 25355, A
C 33	63.5	1.2	67	1	US-08-471-780C-126	Sequence 126, App
C 34	63.5	1.2	67	1	US-08-467-282B-126	Sequence 126, App
C 35	63.5	1.2	67	2	US-08-471-282A-126	Sequence 126, App
C 36	63.5	1.2	67	2	US-08-466-710C-126	Sequence 126, App
C 37	63.5	1.2	67	3	US-08-468-739C-126	Sequence 126, App
C 38	63.5	1.2	67	4	US-09-293-769A-126	Sequence 126, App
C 39	63.5	1.2	67	4	US-08-537-871A-9	Sequence 9, Appl
C 40	63.5	1.2	74	3	US-08-679-493A-33	Sequence 33, Appl
C 41	63	1.2	63	4	US-09-134-000C-6110	Sequence 6130, Ap
C 42	63	1.2	74	4	US-09-270-767-34440	Sequence 34440, A
C 43	63	1.2	74	4	US-09-270-767-49657	Sequence 49657, A
C 44	63	1.2	78	4	US-09-949-016-10753	Sequence 10753, A
C 45	63	1.2	86	4	US-09-252-991A-29383	Sequence 29383, A

ALIGNMENTS

RESULT 1
US-08-967-513-3
; Sequence 3, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: Of Use For Determination
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967, 513
; FILING DATE: 11-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/687, 645
; FILING DATE: July 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20, 931
; REFERENCE/DOCKET NUMBER: MSU 4.1-309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100

TELEFAX: (517) 347-4103
TELEX: No. 5783436e
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 100
TYPE: amino acids
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Protein
DESCRIPTION: Protein
HYPOTHETICAL: No
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: N/A
STRAIN: N/A
INDIVIDUAL ISOLATE: N/A
CELL TYPE: N/A
FEATURE:
NAME/KEY: subunit Urea
LOCATION:
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION: encoded subunit of
US-08-967-513-3

Alignment Scores:
Pred. No.: 8,43e-19 Length: 100
Score: 281.50 Matches: 59
Percent Similarity: 77.00% Conservative: 18
Best Local Similarity: 59.00% Mismatches: 22
Query Match: 5.47% Indels: 1
DB: 1 Gaps: 1

US-09-904-994B-1 (1-2883) x US-08-967-513-3 (1-100)

QY 206 GTGAACTCACACCCCAAGAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCT 265
Db 1 MetGluLeuThrProArgGluLysAspLysLeuLeuLeuPheThrAlaLeuValAla 20
QY 266 AGAAGCGCAAGAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db 21 GluArgArgLeuAlaArgGlyLeuLysLeuAsnTyrProGluSerValAlaLeuIleSer 40
QY 326 GCCCATATTATGAGAGCGCGCGCTGGAAGAAACCGTTCGCCAGCTTATGGAAGAG 385
Db 41 AlaPheIleMetGluGlyAlaArgAspGly--LysSerValAlaSerLeuMetGluGlu 59
QY 386 TGCATGCACTTTTGAAGAAAGATGAATATGCCCGGAGTGGTAATATGTTCCCGAT 445
Db 60 GlyArgHisValLeuThrArgGluGluGlnValMetGluGlyValProGluMetIleProAsp 79
QY 446 CTAGGTGTGAAGCCACCTTTCTGTATGTATGTAACCTGTATCTGTAAATTGGCCCATC 505
Db 80 IleGlnValGluAlaThrPheProAspGlySerLysLeuValThrValHisAsnProIle 99

RESULT 2
US-08-687-645B-3
Sequence 3, Application US/08687645B
Patent No. 5846752

GENERAL INFORMATION:
APPLICANT: Robert P. Hausinger
TITLE OF INVENTION: Mutant Urease and Method
TITLE OF INVENTION: of Use For Determination
TITLE OF INVENTION: of Urea
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,645B
FILING DATE: July 26, 1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 100
TYPE: amino acids
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Protein
DESCRIPTION: Protein
HYPOTHETICAL: No
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: N/A
STRAIN: N/A
INDIVIDUAL ISOLATE: N/A
CELL TYPE: N/A
FEATURE:
NAME/KEY: subunit Urea
LOCATION:
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION: encoded subunit of
US-08-687-645B-3

Alignment Scores:
Pred. No.: 8,43e-19 Length: 100
Score: 281.50 Matches: 59
Percent Similarity: 77.00% Conservative: 18
Best Local Similarity: 59.00% Mismatches: 22
Query Match: 5.47% Indels: 1
DB: 2 Gaps: 1

US-09-904-994B-1 (1-2883) x US-08-687-645B-3 (1-100)

QY 206 GTGAACTCACACCCCAAGAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCT 265
Db 1 MetGluLeuThrProArgGluLysAspLysLeuLeuLeuPheThrAlaLeuValAla 20
QY 266 AGAAGCGCAAGAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db 21 GluArgArgLeuAlaArgGlyLeuLysLeuAsnTyrProGluSerValAlaLeuIleSer 40
QY 326 GCCCATATTATGAGAGCGCGCGCTGGAAGAAACCGTTCGCCAGCTTATGGAAGAG 385
Db 41 AlaPheIleMetGluGlyAlaArgAspGly--LysSerValAlaSerLeuMetGluGlu 59
QY 386 TGCATGCACTTTTGAAGAAAGATGAATATGCCCGGAGTGGTAATATGTTCCCGAT 445
Db 60 GlyArgHisValLeuThrArgGluGluGlnValMetGluGlyValProGluMetIleProAsp 79
QY 446 CTAGGTGTGAAGCCACCTTTCTGTATGTATGTAACCTGTATCTGTAAATTGGCCCATC 505
Db 80 IleGlnValGluAlaThrPheProAspGlySerLysLeuValThrValHisAsnProIle 99

```

RESULT 3
US-07-732-242C-1
; Sequence 1, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi,Takeeshi; Masaki,Haruhiko;
; APPLICANT: Hozumi,Makoto; Nakamura,Akira;
; APPLICANT: Maeda,Michihisa; Yonetsu,Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frieauf,Holtz,Goodman & Woodward,P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 972-1400
; TELEFAX: (212) 370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-732-242C-1

Alignment Scores:
Pred. No.: 1,25e-17 Length: 99
Score: 269.50 Matches: 55
Percent Similarity: 73.74% Conservative: 18
Best Local Similarity: 55.56 Mismatches: 25
Query Match: 5.24% Indels: 1
DB: 1 Gaps: 1

US-09-904-994B-1 (1-2883) x US-07-732-242C-1 (1-99)

Oy      209 AAACCTCACCCCAAGAGCAGAAGAAAGTTCTTGTAATTATTCGGCGGGAAGTGCCTAGA 268
       ||||| | | | | : || | | | | | | | | | | | | | | | | | | | | | | |
Db      1 LysleuthrThSerArglumetguluylsleumetilvalalalaapleualarg 20
Oy      269 AAGCGCAAGAGAGGCGTTAAAGCTCAACCAACCGGAGACCATTGGTTACATTAGTCC 328
       ::||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      21 ArgAsglySgluaArgglyLeuylsleuAenTyxProgluAlaValAlametiIethrTx 40
Oy      329 CATATTATGACGAGAGCGGCGGTGGAAAAAACCGTGGCCAGCTTAGGAAGACTGC 388
       ::::: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      41 GluvallelengluGlyAlaArgSepgly---LysThrValAlaGlnleuemetGlnTyrgly 59
Oy      389 ATGCACTTTGAAAAAAGATGAAGTAATGCCCGGGGGTGGATAATATGTTCCCGATCTA 448
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      60 AlaThrIleuethrtyrSgluaSepValmetGluGlyValAlaGlnleuetIleproapePle 79
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      449 GGtGTGAAGCAACCTTTCCTGATGtgTAcGAAACTTGTAACtGTGAATggccCATc 505

```

```

Db      80 GlnleclunlaThrPheProAspGlyThrIlyrIleuValThrValHisasproIle 98
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
RESULT 4 1
US-08-467-822-23
; Sequence 23, Application US/08467822
; Patent No. 5843460
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauterbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,822
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495_0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-822-23
Alignment Scores:
Pred. No.: 7,61e-17 Length: 100
Score: 261.50 Matches: 56
Percent Similarity: 73.00% Conservative: 17
Best Local Similarity: 56.00% Mismatches: 26
Query Match: 5.08% Indels: 1
Gaps: 1
US-09-904-994B-1 (1-2883) x US-08-467-822-23 (1-100)
Oy      206 GTGAATCTCACACCCAAAGAGCAAGAAAGTTCTTTATATATATGCGGCGAGTGGCT 265
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      1 MecdluethurProrAkrgrlulrYsaeplyrIleuIleuPheThrAlaGlyIleuValHis 20
      266 AGAAAGCGCAAGACAGAGGCGCTTAAGACTCAACCAACCCGAAGCATTGCTTACATTAGT 325

```

Db 21 GIUARGARGLEUALALYSGLYSEULYSLEUAENTYRPROGLUARGVALALALEULESER 40
QY 326 GCCCATATTATGACGAGCGCCCGCTGGAAAAAACCCTTGCCACGCTTATGGAAG 385
Db 41 CYSALALILEMETGLUGLYALARGGLUGLY--LYRTHRVALALAGLNUWETSERGLU 59
QY 386 TGCATGCACTTTTGGAAAAAGATGAAGTAATGCCGGGCGGTGATATGTTCCCGAT 445
Db 60 GLYARGTHRVALLEUHRALAGLUGLNUWETGLUGLYALPROGLUWETILEYSAAP 79
QY 446 CTAGGTGTAAGAACCACTTCTCGATGATGACGAACTGTGAATGGAATGGCCCATC 505
Db 80 VALGINVALGLUCYTHRPEPROAEPGLYTHRLYSLEUVALSERILEHISSEPROILE 99

RESULT 5
US-08-432-697-23
Sequence 23, Application US/08432697
Patent No. 6248330
GENERAL INFORMATION:

APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-432-697-23

Alignment Scores:
Pred. No.: 7.61e-17 Length: 100
Score: 261.50 Matches: 56
Percent Similarity: 73.00% Conservative: 17
Best Local Similarity: 56.00% Mismatches: 26
Query Match: 5.08% Indels: 1
DB: 3 Gaps: 1

US-09-904-994b-1 (1-2883) x US-08-432-697-23 (1-100)

QY 206 GTGAACCTGACACCAAGAGCAAGAAAGTTCTGTATTATTATGCGGCGCAAGTGCGT 265

Db 1 METGLULEUHRTHRPROARGLNUYSAPLYSLEULEUHRPHETHRVALAGLYLEUVALALA 20
QY 266 AAAAAAGCGAAGACGAGGCGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTACT 325
Db 21 GIUARGARGLEUALALYSGLYSEULYSLEUAENTYRPROGLUARGVALALALEULESER 40
QY 326 GCCCATATTATGACGAGCGCCCGCTGGAAAAAACCCTTGCCACGCTTATGGAAG 385
Db 41 CYSALALILEMETGLUGLYALARGGLUGLY--LYRTHRVALALAGLNUWETSERGLU 59
QY 386 TGCATGCACTTTTGGAAAAAGATGAAGTAATGCCGGGCGGTGATATGTTCCCGAT 445
Db 60 GLYARGTHRVALLEUHRALAGLUGLNUWETGLUGLYALPROGLUWETILEYSAAP 79
QY 446 CTAGGTGTAAGAACCACTTCTCGATGATGACGAACTGTGAATGGAATGGCCCATC 505
Db 80 VALGINVALGLUCYTHRPEPROAEPGLYTHRLYSLEUVALSERILEHISSEPROILE 99

RESULT 6
US-08-466-248-23
Sequence 23, Application US/08466248
Patent No. 6258359
GENERAL INFORMATION:

APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,248
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-248-23

Alignment Scores:
Pred. No.: 7.61e-17 Length: 100
Score: 261.50 Matches: 56
Percent Similarity: 73.00% Conservative: 17
Best Local Similarity: 56.00% Mismatches: 26
Query Match: 5.08% Indels: 1
DB: 3 Gaps: 1

US-09-904-994B-1 (1-2883) x US-08-466-248-23 (1-100)

QY 206 GTGAACTCACACCCAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCGT 265
Db 1 MecJleuthrProAargGluYAspIysleuLeuLeuPheTrAlaGlyLeuValAla 20

QY 266 AGAAAGCGCAAGAGAGCGCTTAAGCTCAACCAACCGGAGCCATTGCTTACATTAGT 325
Db 21 GluArgArgLeuAlaIalysGlyLeuLeuLeuValTyrProGluArgValAlaLeuIleSer 40

QY 326 GCCCATTTTATGACGAAAGCGCGCGTGGAAAAAACCCTTGGCCAGCTTATGGAAGAG 385
Db 41 CysAlaIleMetGluGlyAlaArgGluGly---LysThrValAlaGlnLeuMetSerGlu 59

QY 386 TGCATGACCTTTTGTAAAAAAGATGATGCGCGGCGGTGAATATGTTCCCGAT 445
Db 60 GluArgThrValIleuThrAlaGlnValMetGluGlyValProGluMetIleLysAsp 79

QY 446 CTAGGTGTAGAAAGCCACTTCTGTATGTATGCAAACTTGTACTGTGAATTGGCCCATC 505
Db 80 ValGlnValGluCysThrPheProAspGlyThrIlysLeuValSerIleHisSerProIle 99

RESULT 7
US-09-602-777A-18
Sequence 18, Application US/09602777A
Patent No. 6831165
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schröder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
FILE REFERENCE: BGI-128CP
CURRENT APPLICATION NUMBER: US/09/602, 777A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932127.2
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932129.9
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: DE 19932226.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932920.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932924.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932928.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932930.3
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932933.8

PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932935.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932973.7
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933002.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933003.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941390.8
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941391.6
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 442
SEQ ID NO 18
LENGTH: 100
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-602-777A-18

Alignment Scores:
Pred. No.: 9.05e-16 Length: 100
Score: 250.50 Matches: 51
Percent Similarity: 72.45% Conservative: 20
Best Local Similarity: 52.04% Mismatches: 26
Query Match: 4.87% Indels: 1
DB: 4 Gaps: 1

US-09-904-994B-1 (1-2883) x US-09-602-777A-18 (1-100)

QY 212 CTCACACCCAAAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCTAGAAG 271
Db 3 IleThrProArgGluGlnGluIlyLeuMetIleValValAlaIleAspLeuAlaArg 22

QY 272 CGCAAGCAGAGGCGCTTAAGCTCAACCAACCGGAGCCATTGCTTACATTAGTCCCAT 331
Db 23 ArgIysAspArgGlyLeuIlyLeuAsnHisProGluAlaValAlaLeuIleThrGlu 42

QY 332 ATTTATGACGAGAGCGCGCTTGAAGAAAAAACCCTTCCAGCTTATGGAAGTGCATG 391
Db 43 LeuIleGluGlyAlaArgAspGly--ArgThrValAlaAspLeuMetSerTyrGlySer 61

QY 392 CACTTTTGTAAAAAAGATGAGTATGCCCGGCGGTGAATATGTTCCGATCTAGT 451
Db 62 ThrIleuThrArgAspAspValLeuGluGlyIleProGluMetIleProAspIleGln 81

QY 452 GTAGAGCCACCTTCTGTATGTATGCAAACTTGTAACTGTGAATTGGCCCATC 505
Db 82 ValGlnIleThrPheAspAspGlyThrIlyLeuValThrValHisAsnProIle 99

RESULT 8
US-09-252-991A-17446
Sequence 17446, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252, 991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074, 788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 17446
 LENGTH: 92
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17446

Alignment Scores:

Pred. No.:	7.43	Length:	92
Score:	87.50	Matches:	36
Best Similarity:	37.76%	Conservative:	18
Best Local Similarity:	25.17%	Mismatches:	24
Query Match:	1.70%	Indels:	65
DB:	4	Gaps:	7

US-09-904-994b-1 (1-2883) x US-09-252-991A-17446 (1-92)

```

QY 2101 AGCGCTACATCTCCCAATACATCAACCCGCTTGAACCAACGCGGTGAGGAGTATA 2160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6 SerAlaThrSerProSerLeuSerAlaProProAla-----AlaSerSer 20
QY 2161 TCGGCTCTGTGAGAGGAGGCAAGATCGCGACTTGTGTGTGGAATCCTGCTTTTGG 2220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 21 SerSerSerTriplysArgSerAla----- 28
QY 2221 GCGTAAACCCAAATCGTGATCAAAAGCGGTATGTGCTTCTCTGAAATGCGGAT 2280
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 29 -----AlaAlaAlaArgTyrSerVal 35
QY 2281 CTAACGCTGTGCGCCACTCCCAACGCGTTTATTACGCGAAATGTTGGGCATCAG 2340
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 36 -----ThrArgCysMet----- 39
QY 2341 GCAAGCGAAATTGACACACGATCATTTGTTCAAAGTGGCTATGAAATGGCG 2400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 40 -----SerAlaIleSerCysSerSerProAlaIlySerAla 51
QY 2401 TGAAG-----AAAGCTTGGGCTTGAAGCGCAAGTTCTACCGGTCAAAACTGCC 2451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52 AlaIlyArgArgProArgSerTyrArg-----ArgSerCysThrSer 65
QY 2452 GTAACATCAACGAAGAAGCTTCAGAGTTCACAGCAAAACGCAAAATCACCCTGATC 2511
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 66 AlaSerSerProGlyThrSerAlaLeuTyr-----ArgArgAlaSerProIlyle 83
QY 2512 CGAAAACCT 2520
    ||||| |||||
DB 84 ArgValPro 86
  
```

RESULT 9

US-09-328-352-4313
 Sequence 4313, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 4313
 LENGTH: 96
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-4313

Alignment Scores:

Pred. No.:	57.8	Length:	96
Score:	78.50	Matches:	22
Percent Similarity:	41.35%	Conservative:	21

Best Local Similarity:	21.15%	Mismatches:	32
Query Match:	1.53%	Indels:	29
DB:	4	Gaps:	3

US-09-904-994b-1 (1-2883) x US-09-328-352-4313 (1-96)

```

QY 1381 GACCTGTGCGCCCTCCAAACATGTTGTGAACCCATTGGCTAGACCGGTAGGAATTGT 1322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 9 GlyAspValProProPro----- 14
QY 1321 TGTGAGAAAGAAAGTGGGTGTGTGAATCGATTCCCCAGCGGTATATCATCTTCC 1262
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 15 -----IleMetProProValThrProProIleIleSerValThrSer 28
QY 1261 CCTGCTAGTCTTCTGTGCGCACACCGACAGCATATAGAGGCTTACGCACTTGACG 1202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 29 ProArgGlyLeuSerProProThrProPro-----ValThrProProIleIle 44
QY 1201 TCTTGTTCCTCCCTTGCACATGCAATGATTTGCGGTTTATCCCAATGTGCGCT 1142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 45 SerValIleThrSerSerGlyLeuProSerThrProProValThrProProIleThrSer 64
QY 1141 TTGTAAATCCCGGTACTGATA-----ATCATCGCTTACTGATGACT 1097
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 ValThrLeuProAlaThrProValThrIleProValAlaValProPheValIleSer 84
QY 1096 AAATCTAGGCTG 1085
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 85 ValSerArgIle 88
  
```

RESULT 10

US-08-928-081-1
 Sequence 1, Application US/08928081
 Patent No. 5985631
 GENERAL INFORMATION:
 APPLICANT: Soman, Gopalan
 APPLICANT: Thomas, Jr., William D.
 APPLICANT: Monath, Thomas P.
 TITLE OF INVENTION: Stabilization of
 TITLE OF INVENTION: Helicobacter Urease
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: Clark & Ebling LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,081
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 06132/023001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-428-0200
 TELEFAX: 617-428-7045
 TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-928-081-1

Alignment Scores:

Pred. No.: 27 Length: 15
Score: 77.00 Matches: 13
Percent Similarity: 93.33% Conservative: 1
Best Local Similarity: 86.67% Mismatches: 1
Query Match: 1.50% Indels: 0
DB: 2 Gaps: 0

US-09-904-994B-1 (1-2883) x US-08-928-081-1 (1-15)

QY 1527 GAAGCGGCGCGATTGTTTAAATTCATGATGAGACTGGGCGACA 1571

Db 1 GluAaGlyAlaIleGlyPheAlaIleHisGluSerTrpGlyThr 15

RESULT 11

US-09-338-920B-1
Sequence 1, Application US/09338920B
Patent No. 6709851
GENERAL INFORMATION:

APPLICANT: Soman, Gopalan

APPLICANT: Thomas, William D., Jr.

APPLICANT: Monath, Thomas P.

TITLE OF INVENTION: Stabilization of Helicobacter Urease

FILE REFERENCE: 06132/023002

CURRENT APPLICATION NUMBER: US/09/338,920B

PRIOR FILING DATE: 2002-12-23, 081

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 15

TYPE: PRT

ORGANISM: Helicobacter pylori

US-09-904-994B-1 (1-2883) x US-09-338-920B-1 (1-15)

QY 1527 GAAGCGGCGCGATTGTTTAAATTCATGATGAGACTGGGCGACA 1571

Db 1 GluAaGlyAlaIleGlyPheAlaIleHisGluSerTrpGlyThr 15

RESULT 12

US-09-489-039A-7356
Sequence 7356, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7356

LENGTH: 93

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

Alignment Scores:

Pred. No.: 274 Length: 93

Score: 71.50 Matches: 32
Percent Similarity: 37.27% Conservative: 9
Best Local Similarity: 29.09% Mismatches: 46
Query Match: 1.39% Indels: 23
DB: 4 Gaps: 6

US-09-904-994B-1 (1-2883) x US-09-489-039A-7356 (1-93)

QY 2026 TGCCTAAGTTCGAGCAATCCTGCGCTTCA-----CGCCCC-----ATTGCTTGC 1982

Db 2 CysArgSerThrSerAlaValAlaProCysSerThrIleArgProAlaIleThr 21

QY 1981 GAATCCAGCTTTCATGCGATGACACCATATATGAGACATCTTCAGCCGCGATA 1922

Db 22 ThrCysSerAlaIleCysAlaIleThrProArgSerTrpValIleSerIleIleAla 41

QY 1921 GAGCCGGGCGGATACGCTTTTGAGAAATGTGAAATCCTGCGGATGCGTTGCTACG 1862

Db 42 TrpLeuAlaArgAlaArg-----SerArgSerArgArgArgIle----- 54

QY 1861 TCGTGCATGTCATGACATGTCATGCTTTCGAAACCGTATTAAGTATAGGATA 1802

Db 55 -----CysAlaCysThrValThrSerSerAlaValAlaGlySer 67

QY 1801 GTGGGGGTGAGGAGAGGAGGATATATGAGCTGCGCTGCATGATGATGATGATGAT 1742

Db 68 SerAlaIleSerSerLeuGlyArgGlnSerAlaMetAlaIleIleThr----- 84

QY 1741 GAGTGTCTCCACCGCTCCCTCAATGTGG 1712

Db 85 ArgCysSerMetPro---ProGluSerTrp 93

RESULT 13

US-09-489-039A-12832
Sequence 12832, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 12832

LENGTH: 89

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12832

Alignment Scores:

Pred. No.: 374 Length: 89
Score: 70.00 Matches: 27
Percent Similarity: 46.07% Conservative: 14
Best Local Similarity: 30.34% Mismatches: 22
Query Match: 1.36% Indels: 26
DB: 4 Gaps: 5

US-09-904-994B-1 (1-2883) x US-09-489-039A-12832 (1-89)

QY 645 TGGAAATGCTACCCACATGCAAGATTAGGC---CCTTCATTAGTAACCTCAAGTTGC 589

Db 2 TrpIleThrIleAlaIleThrProArgArgAlaGlyThrThrSerPheIleThrSerPro 21

QY 588 GTTACTCTTTCGCTGATGAGCTGATGCTTTTATGCAACCAATTCACCTTCGCGCC 529

Db 22 IleSerThrSerProSerLeuThrSer-----SerSerPro 33

QY 528 GCTTTGAAGTCTCATCTGCTGATGAGGCAATTCACAGTTCAAGTTTCGATCACTCA 469

Db 34 AlaIleMetArgSerSerValAlaPheProGln-----ProGlu 46

```
QY 468 GGA-----AAGGTGGCTTCTACACTATA-----TCGGGA 439
Db 47 GlyProThrLysThrIleAsnSerSerIleSerArgLeuThrTrpCysSerVal 66
QY 438 ACCATATACCAACCCCGGCGCATTTACT 412
Db 67 ThrValLeuProAsnLeuLeuValThr 75

RESULT 14
US-09-328-352-4788
; Sequence 4788, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4788
; LENGTH: 99
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4788

Alignment Scores:
Pred. No.: 499 Length: 99
Score: 69.00 Matches: 33
Percent Similarity: 32.65% Conservative: 15
Best Local Similarity: 22.45% Mismatches: 37
Query Match: 1.34% Indels: 62
DB: 4 Gaps: 8

US-09-904-994B-1 (1-2883) x US-09-328-352-4788 (1-99)
QY 1573 GTTGGCCCCAGCTTCATGACATTTAAACCAATGCGCGCGCTTCTACTTGTCTACA 1514
Db 1 lleileProvalSer-----GlyLeuSerProArgSerPro----- 12
QY 1513 AGTTGTTTTTGTAGAGTATTGCTTTGCCCAAAAGCCCATTCATGAGATACGT 1454
Db 13 -----PrometProPrometProProilleMetPro----- 22
QY 1453 TCTGCTCGCGCAACATGCGGTGCAAGTTCATTTGCCGAGATGATAGTCGATTC 1394
Db 23 -----ValIleGlyLeuLysIle 28
QY 1393 GTGCCATCTACAGACCTGCGCGCT-----CCA 1364
Db 29 SerProSerPro-----ProIleProPrometIleProvalIleGlyLeuLysIleSerPro 47
QY 1363 AACATGTTGTAAACGCGCATTTGCTAGAGCGTAGAGGAATTTGTTGAGAGAAAGAGTGG 1304
Db 48 SerProProIleProProIleThrProValIleGly----- 59
QY 1303 GTGTGTGAATCCATTTCCCGAGGGTAATATATACCTTCCCTTCGTAGTGTCTGTG 1244
Db 60 -----ProLysIleSerProvalProSerProProIleAsnLys 73
QY 1243 CCCAGCACCGACGACATGAGGGCTT-----ACGCGATCTTGATGTCCTTG 1196
Db 74 GluLeuProAsnThrLeuIleGlyValIleGlyAlaThrThrSerThrSerPro 93
QY 1195 TTTCCTGCTTGGCAATGCCA 1175
Db 94 LeuPro---LeuProMetPro 99

RESULT 15
US-09-602-777A-12
; Sequence 12, Application US/09602777A
; Patent No. 6831165
```

```
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 12
; LENGTH: 40
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-12

Alignment Scores:
Pred. No.: 327 Length: 40
Score: 68.50 Matches: 14
Percent Similarity: 48.89% Conservative: 8
Best Local Similarity: 31.11% Mismatches: 16
```


Query Match: 1.33% Indels: 7
DB: 4 Gaps: 1

US-09-904-994B-1 (1-2883) x US-09-602-777A-12 (1-40)

```
QY      509 CCAGATGAGCACTTCAAAGCGGCGAAGTGAATTGGTTGCGATTAAGACATGAGCTC 568
      ||| |||::: ||| |||
Db      3 ProGlyGluTyrIleLeuSerSerGluSerLeuThrGly----- 15
QY      569 AATGCAGGCAAGAGTAACCGAACTTGAGGTACTAATGAGGCGCTAAATCCTTGCA 628
      ||| |||::: ||| |||
Db      16 AsnValGlyArgGluAlaLysThrIleGluIleAsnThrGlyAspArgProValGln 35
QY      629 GTGGTAGCCATTTC 643
      ||| ||| ||| ||| |||
Db      36 IleGlySerHisPhe 40
```

Search completed: February 15, 2005, 20:34:27
Job time : 36.5 secs

This Page Blank (uspto)

This Page Blank (uspto)